

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Analyse de la complexité polynomiale de deux méthodes de points intérieurs basées sur la trajectoire centrale, en programmation linéaire

LESEULTRE, Françoise

Award date:
1992

Awarding institution:
Université de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**FACULTÉS UNIVERSITAIRES NOTRE-DAME DE LA PAIX
NAMUR
FACULTÉ DES SCIENCES**

**Analyse de la complexité polynomiale
de deux méthodes de points intérieurs,
basées sur la trajectoire centrale,
en programmation linéaire.**

Promoteur : Monsieur le Professeur J-J Strodiot

Françoise Leseultre

Année académique : 1991- 1992

Je tiens à remercier très sincèrement Monsieur le Professeur Jean-Jacques Strodiot pour ses conseils, sa patience et sa disponibilité qui m'ont permis de réaliser le présent travail.

J'exprime également ma gratitude à ma famille et à mes amis pour leur soutien tout au long de mes études.

En présentant ce travail, mon but et mon espoir sont de montrer que je n'ai pas démerité de ceux qui me consacreront leur temps et leur connaissance.

Table des matières

Introduction

4

Chapitre I La complexité d'un algorithme

7

1. Avant-propos
2. La complexité d'un algorithme
3. Encodage de nombres
4. La complexité polynomiale d'un algorithme
 - 4.1. Définitions
 - 4.2. Quelques caractéristiques
5. L'algorithme du simplexe: complexité exponentielle

7

7

7

16

16

17

18

Chapitre II Deux méthodes de minimisation en programmation non linéaire

30

1. La méthode "barrière"
 - 1.1. Présentation de la méthode
 - 1.2. Les fonctions barrières
 - 1.3. Convergence de la méthode
 - 1.4. Modification de la méthode
2. La méthode S-S-D
 - 2.1. La transformation "scaling"
 - 2.2. Méthode de "plus forte pente"
 - 2.3. Utilité du scaling dans la méthode
 - 2.4. L'algorithme S-S-D

30

30

31

34

38

40

40

41

43

45

Chapitre III La trajectoire centrale

46

1. Le problème de programmation linéaire
 - 1.1. Présentation
 - 1.2. Le concept de projection
 - 1.3. Le problème dual
2. La fonction barrière logarithmique
 - 2.1. Définition et propriétés immédiates
 - 2.2. Variation de la fonction barrière autour du point e
 - 2.3. Effet du " scaling " sur la fonction barrière logarithmique
3. Les fonctions auxiliaires
 - 3.1. La fonction de pénalité intérieure
 - 3.2. La fonction potentielle
4. La trajectoire centrale
 - 4.1. Définition de la trajectoire centrale
 - 4.2. Les paramétrages de la trajectoire centrale

46

46

47

48

50

50

50

55

56

56

59

69

69

71

4.2.1. Paramétrage primal-dual	71
4.2.2. Paramétrage dual	73
4.3. Identification des trajectoires caractérisées par les deux paramétrages	73
Chapitre IV De l'algorithme conceptuel à l'algorithme pouvant être mis en oeuvre	78
1. L'algorithme conceptuel	78
2. L'algorithme de minimisation interne	79
2.1. Le critère de proximité	80
3. Mise à jour du paramètre	83
4. Critère d'arrêt de l'algorithme principal	85
5. L'algorithme pouvant être mis en oeuvre	87
5.1. Complexité de l'algorithme	88
Chapitre V La méthode de la fonction de pénalité intérieure	90
1. Préliminaires	90
2. L'algorithme de minimisation interne	91
2.1. La direction Newton-Raphson	91
2.2. Coïncidence des directions S-S-D et Newton-Raphson	95
3. Le critère de proximité	97
4. L'algorithme pouvant être mis en oeuvre	106
4.1. Le critère d'arrêt l'algorithme principal	107
4.2. L'algorithme principal	108
5. Etude de la complexité de l'algorithme	109
Chapitre VI La méthode de la fonction potentielle	113
1. Préliminaires	113
2. La direction S-S-D	113
3. Le critère de proximité et la mise à jour du paramètre v	114
4. Un algorithme de complexité $O(nL)$	120
4.1. L'algorithme	120
4.2. Etude de la complexité	122
5. Un algorithme de complexité $O(\sqrt{n} L)$	130
5.1. L'algorithme	130
5.2. Etude de la complexité	132
Conclusion	144
Bibliographie	145

Introduction

Dans ce mémoire, nous nous intéressons à la résolution du problème de *programmation linéaire* suivant :

$$(PL) \quad \left[\begin{array}{ll} \text{Minimiser} & c^T x \\ \text{sous} & A x = b \\ & x \geq 0 \end{array} \right.$$

où $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, A est une matrice de dimension (m,n) avec $n > m$ et est de rang plein.

Nous supposons que la région admissible

$$S = \{x \in \mathbb{R}^n \mid A x = b, x \geq 0\}$$

est bornée et a un intérieur relatif non vide défini par

$$\overset{\circ}{S} = \{x \in \mathbb{R}^n \mid A x = b, x > 0\}.$$

La méthode du simplexe développée par Dantzig s'est avérée très efficace pour résoudre un tel problème. Cependant, sa complexité exponentielle la rend moins désirable par rapport aux algorithmes de *complexité polynomiale*. Cette propriété sera étudiée dans le premier chapitre.

Nous nous tournerons ainsi vers des algorithmes de type Karmarkar [7].

A la différence de l'algorithme du simplexe, ceux-ci se développent à l'intérieur de la région admissible en suivant une trajectoire particulière.

Cette trajectoire, le long de laquelle le coût diminue, est appelée la *trajectoire centrale*. Elle peut être interprétée dans le cadre des méthodes suivantes [4] :

- Méthode des centres.
- Méthode "primal-dual".
- Méthode de la fonction de pénalité intérieure.
- Méthode de la fonction potentielle.

Nous nous attacherons à la description des deux dernières méthodes. Leur algorithme repose sur deux méthodes de programmation non linéaire, à savoir la méthode barrière et la méthode S-S-D, décrites dans le second chapitre.

Nous consacrerons le troisième chapitre à l'étude des propriétés des fonctions de pénalité intérieure et potentielle, toutes deux construites à partir de la fonction barrière logarithmique.

Ces fonctions donnent lieu à un "paramétrage" différent de la trajectoire centrale qui comprend l'ensemble de leurs points optimaux et qui converge vers une solution optimale du problème (PL).

Un algorithme unifié sera dans un premier temps construit dans le chapitre IV pour être ensuite particularisé aux deux fonctions auxiliaires dans les deux derniers chapitres. Nous obtiendrons pour chacune des algorithmes de complexité polynomiale : selon les hypothèses initiales, $O(nL)$ ou $O(\sqrt{n} L)$ itérations où n est le nombre de variables et L est la longueur des entrées du problème.

Pour l'étude des méthodes de fonction de pénalité intérieure et potentielle, nous nous sommes référés aux articles [2] et [3] de C.Gonzaga.

NOTATION

Nous dressons ci-dessous la liste des principaux symboles utilisés dans le texte :

- $\mathbb{Q}, \mathbb{N}, \mathbb{Z}$: sans autre précision, ce sont les ensembles rationnels, naturels et entiers.
- \mathbb{R} : l'espace réel.
- $\mathbb{R}_+^n, \mathbb{R}_{++}^n$: ensembles des vecteurs non négatifs (≥ 0) et positifs (> 0) de \mathbb{R}^n .
- $e = [1 \dots 1]^T$.
- $X = \text{diag}(x_1, \dots, x_n)$.
- x, w, z, A, b, c : variables du problème primal et dual.
- $y, \bar{w}, \bar{z}, \bar{A}, \bar{b}, \bar{c}$: variables des problèmes dans l'espace transformé par un "scaling".
- $\hat{x}, \hat{w}, \hat{z}$: solutions optimales pour les problèmes primal et dual.
- $\hat{v} = c^T \hat{x}$: valeur optimale.
- $S, \overset{\circ}{S}$: ensemble admissible pour (PL) et son intérieur relatif.
- $Z, \overset{\circ}{Z}$: ensemble admissible du problème dual et son intérieur relatif.
- P_M, \bar{P}_M : matrice de projection sur le noyau de M et son complémentaire orthogonal.
- $p(\cdot)$: fonction barrière.
- $f_\alpha(\cdot), f_q(\cdot)$: fonction de pénalité intérieure et fonction potentielle.
- $\rho \rightarrow x(\rho)$: paramétrage générique de la trajectoire centrale.
- α : paramètre de pénalité.
- v : paramètre de la fonction potentielle, borne inférieure du coût.
- q : constante fixée telle que $q \geq n$.
- $h(x, \rho)$: direction S-S-D issue de x .

- $\bar{h}(e, \rho) = X^{-1} h(x, \rho)$: direction S-S-D issue de e dans l'espace transformé par un " scaling ".
- $\delta(x, \rho) = \|\bar{h}(x, \rho)\|$: mesure de proximité de x par rapport à $x(\rho)$.
- $\Delta(\rho)$: saut de dualité associé au paramètre ρ .
- \log : sans autre précision (ex.: $\log_2 x$), est la fonction logarithme en base e .

Chapitre I

La complexité d'un algorithme

1. Avant-propos

Ce chapitre présente quelques notions de base relatives à la théorie de complexité, [5] et [10]. Nous aborderons de manière plus précise la complexité polynomiale d'un algorithme en décrivant ses rapports avec la complexité exponentielle, illustrée par l'algorithme du simplexe.

2. La complexité d'un algorithme

Le coefficient de performance le plus important pour un algorithme est le temps d'exécution. Nous exprimons les exigences de temps d'un algorithme en termes de nombre de pas, d'étapes, d'opérations arithmétiques effectuées.

Ainsi, nous mesurons la complexité d'un algorithme par l'ordre de grandeur du nombre d'opérations nécessaires pour trouver une solution exacte au problème et ce, en se plaçant dans la situation la plus défavorable.

Le nombre d'opérations est calculé en fonction de la taille des entrées du problème. Le concept de "taille" peut être formulé en considérant un schéma d'encodage qui représente les données d'un algorithme comme une suite de symboles.

Nous définissons la taille de l'entrée comme la longueur de cette suite, c'est-à-dire le nombre de symboles de la suite.

Différents schémas d'encodage peuvent donner lieu à différentes longueurs d'encodage. En ce qui nous concerne, nous choisissons l'encodage binaire usuel. De cette façon, tout problème est encodé en une suite finie de 0 et de 1.

3. Encodage de nombres

- Pour encoder un entier $n \neq 0$, nous avons besoin d'une cellule pour le signe et de

$$\lceil \log_2 (|n| + 1) \rceil^1$$

cellules pour la suite $\{0, 1\}$ de la représentation binaire de sa valeur absolue. Pour 0, une seule cellule est nécessaire. Donc, l'espace nécessaire pour encoder un entier est:

$$\langle n \rangle = 1 + \lceil \log_2 (|n| + 1) \rceil, \quad n \in \mathbb{Z}$$

Nous appelons $\langle n \rangle$ la longueur de l'encodage de n .

¹ " $\lceil x \rceil$ ", signifie : le plus petit entier supérieur ou égal à x .

Justification du terme $\lceil \log_2 (|n| + 1) \rceil$

$$\forall n \in \mathbb{N}_0, \exists z \in \mathbb{N} : 2^z \leq n < 2^{z+1}$$

La représentation binaire de n demande donc $z + 1$ cellules. Or, il découle de cette relation que

$$2^z < n + 1 \leq 2^{z+1}$$

ce qui est équivalent à

$$z < \log_2 (n + 1) \leq z + 1.$$

Nous avons donc

$$\lceil \log_2 (n + 1) \rceil = z + 1.$$

Ceci explique que la représentation binaire de $|n|$ où $n \in \mathbb{Z}$ demande $\lceil \log_2 (|n| + 1) \rceil$ cellules.

- Chaque nombre rationnel r peut être écrit de manière unique comme $\frac{p}{q}$ où

$p, q \in \mathbb{Z}, q > 0$; p et q sont premiers entre eux.

De cette façon, l'espace nécessaire pour encoder r est :

$$\langle r \rangle = \langle p \rangle + \langle q \rangle$$

$\langle r \rangle$ est appelé la longueur d'encodage de r .

- De manière similaire, la longueur d'encodage d'un vecteur rationnel $\langle x \rangle$ ou d'une matrice $\langle A \rangle$ est définie comme la somme des longueurs d'encodage de ses composantes.

Si $a^T x \leq b$ est une inégalité avec $a \in \mathbb{Q}^n$ et $b \in \mathbb{Q}$, nous disons que $\langle a \rangle + \langle b \rangle$ est la longueur d'encodage de l'inégalité et la longueur d'encodage d'un système d'inégalités est la somme des longueurs d'encodage de ces inégalités.

Pour la commodité de l'écriture, nous écrirons $\langle a_1 \dots a_n \rangle$ pour noter la somme des longueurs d'encodage de toute suite $a_1 \dots a_n$ de matrices ou de vecteurs.

Nous présentons ci-après deux propositions donnant des relations utiles entre les longueurs d'encodage et d'autres fonctions de vecteurs et de matrices:

Proposition 3.1.

a. Pour tout nombre rationnel r , nous avons

$$|r| \leq 2^{\langle r \rangle - 1} - 1. \quad (1)$$

b. Pour tout vecteur $x \in \mathbb{Q}^n$, nous avons

$$\|x\| \leq \|x\|_1 \leq 2^{\langle x \rangle - n} - 1. \quad (2)$$

c. Pour toute matrice $D \in \mathbb{Q}^{n \times n}$, nous avons

$$|\det D| \leq 2^{\langle D \rangle - n^2} - 1. \quad (3)$$

Démonstration

a. Soit $r \in \mathbb{Q}$, nous devons montrer que $|r| \leq 2^{\langle r \rangle - 1} - 1$.

Ceci est équivalent à $\log_2(|r| + 1) \leq \langle r \rangle - 1$ ou encore à $\log_2(|r| + 1) + 1 \leq \langle r \rangle$.

Comme r est un rationnel, il existe deux entiers p et q , premiers entre eux, avec

$q > 0$ tels que $r = \frac{p}{q}$ de façon unique.

De la définition de la longueur de l'encodage d'un nombre rationnel et d'un nombre entier, nous avons:

$$\begin{aligned} \langle r \rangle &= \langle p \rangle + \langle q \rangle \\ &= 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(|q| + 1) \rceil \end{aligned}$$

(Nous ne considérons qu'une seule fois le signe de r).

La thèse s'écrit à présent:

$$\log_2\left(\left|\frac{p}{q}\right| + 1\right) + 1 \leq 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(|q| + 1) \rceil.$$

Nous savons que:

i. $\log_2(|p/q| + 1) \leq \lceil \log_2(|p/q| + 1) \rceil$

ii. $\lceil \log_2(|p| + 1) \rceil + \lceil \log_2(|q| + 1) \rceil \leq \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(|q| + 1) \rceil$

De plus nous pouvons montrer que

$$\lceil \log_2(|p/q| + 1) \rceil \leq \lceil \log_2(|p| + 1) + \log_2(|q| + 1) \rceil \quad (4)$$

En effet, nous avons les équivalences suivantes :

$$\begin{aligned} \lceil \log_2(|p/q| + 1) \rceil &\leq \lceil \log_2(|p| + 1) + \log_2(|q| + 1) \rceil \\ \Leftrightarrow \log_2(|p/q| + 1) &\leq \log_2[(|p| + 1) (|q| + 1)] \end{aligned}$$

$$\Leftrightarrow \left| \frac{p}{q} \right| + 1 \leq (|p| + 1) (|q| + 1)$$

(caractère bijectif de la fonction logarithme)

$$\Leftrightarrow \left| \frac{p}{q} \right| \leq |p| |q| + |p| + |q|$$

$$\Leftrightarrow |p| \leq |p| |q|^2 + |p| |q| + |q|^2$$

Cette dernière inégalité est évidente puisque ses termes sont des entiers positifs avec $q \neq 0$, donc (4) est correct. Ainsi, nous obtenons consécutivement

$$\begin{aligned} \log_2(|p/q| + 1) &\leq \lceil \log_2(|p/q| + 1) \rceil && \text{par i.} \\ &\leq \lceil \log_2(|p| + 1) + \log_2(|q| + 1) \rceil && \text{par (4)} \\ &\leq \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(|q| + 1) \rceil && \text{par ii.} \end{aligned}$$

Par conséquent (1) est prouvé.

- b. Soit $x \in \mathbb{Q}^n$. La première inégalité est une relation bien connue entre les normes 1 et 2.

Avant de montrer la seconde inégalité, nous prouvons le résultat auxiliaire suivant:

$$1 + \sum_{i=1}^n |x_i| \leq \prod_{i=1}^n (1 + |x_i|) \quad (5)$$

dont la preuve se fait par récurrence.

Cette inégalité est immédiate pour $i = 1$.

Supposons ensuite qu'elle est vraie pour $i = n - 1$ et montrons-la pour $i = n$.

Un calcul élémentaire nous donne

$$\begin{aligned}
1 + \sum_{i=1}^n |x_i| &= 1 + \sum_{i=1}^{n-1} |x_i| + |x_n| \\
&= (1 + \sum_{i=1}^{n-1} |x_i|) (1 + |x_n| / (1 + \sum_{i=1}^{n-1} |x_i|)) \\
&\leq (1 + \sum_{i=1}^{n-1} |x_i|) (1 + |x_n|) \\
&\leq \left(\prod_{i=1}^{n-1} (1 + |x_i|) \right) (1 + |x_n|)
\end{aligned}$$

(par hypothèse de récurrence)

$$= \prod_{i=1}^n (1 + |x_i|).$$

Par conséquent l'inégalité (5) est vraie pour $i=n$.
 Venons-en à la preuve du point b. de la proposition:
 Nous avons successivement

$$\begin{aligned}
\|x\|_1 + 1 &= 1 + \sum_{i=1}^n |x_i| \leq \prod_{i=1}^n (1 + |x_i|) \quad \text{par (5)} \\
&\leq \prod_{i=1}^n (2^{\langle x_i \rangle - 1}) \quad \text{par (1)} \\
&= 2^{\sum_{i=1}^n \langle x_i \rangle - n} \\
&= 2^{\langle x \rangle - n}.
\end{aligned}$$

C'est-à-dire la thèse.

- c. Soit $D \in \mathbb{Q}^{n \times n}$. Appelons d_1, \dots, d_n les lignes de la matrice.
 En vertu de l'inégalité d'Hadamard, nous avons

$$|\det D| \leq \prod_{i=1}^n \|d_i\|.$$

Dès lors, nous avons les inégalités suivantes:

$$\begin{aligned}
1 + |\det D| &\leq 1 + \prod_{i=1}^n \|d_i\| \\
&\leq \prod_{i=1}^n (1 + \|d_i\|) \\
&\leq \prod_{i=1}^n 2^{\langle d_i \rangle - n} \quad \text{par (2)} \\
&= 2^{\sum_{i=1}^n \langle d_i \rangle - n^2} \\
&= 2^{\langle D \rangle - n^2}.
\end{aligned}$$

L'inégalité (3) est donc bien démontrée. ■

Proposition 3.2.

$$a. \quad \forall r, s \in \mathbb{Q} \quad \langle rs \rangle \leq \langle r \rangle + \langle s \rangle \quad (6)$$

$$b. \quad \forall D \in \mathbb{Z}^{n \times n} \quad \langle \det D \rangle \leq \langle D \rangle - n^2 + 1 \quad (7)$$

$$c. \quad \forall D \in \mathbb{Q}^{n \times n} \quad \langle \det D \rangle \leq 2 \langle D \rangle - n^2 \quad (8)$$

Démonstration

a. i. Considérons d'abord le cas où r et s sont des entiers. Nous devons montrer que

$$\langle rs \rangle \leq \langle r \rangle + \langle s \rangle,$$

c'est-à-dire que

$$1 + \lceil \log_2(|rs| + 1) \rceil \leq 2 + \lceil \log_2(|r| + 1) \rceil + \lceil \log_2(|s| + 1) \rceil. \quad (9)$$

En utilisant les propriétés de la fonction logarithmique et de la fonction "partie entière", nous pouvons déduire les inégalités suivantes

$$\begin{aligned}
\log_2(|rs| + 1) &\leq \log_2((|r| + 1)(|s| + 1)) \\
\lceil \log_2(|rs| + 1) \rceil &\leq \lceil \log_2(|r| + 1) + \log_2(|s| + 1) \rceil \\
&\leq \lceil \log_2(|r| + 1) \rceil + \lceil \log_2(|s| + 1) \rceil.
\end{aligned}$$

Dés lors, nous avons

$$1 + \lceil \log_2(|rs| + 1) \rceil \leq 2 + \lceil \log_2(|r| + 1) \rceil + \lceil \log_2(|s| + 1) \rceil.$$

C'est-à-dire la thèse (9).

ii. A présent, si r et s sont des nombres rationnels, nous pouvons affirmer qu'il existe p ,

$$q, t, u \text{ entiers tels que } r = \frac{p}{q} \quad (q > 0)^1 \text{ et } s = \frac{t}{u} \quad (u > 0)^2.$$

Il suit alors de la définition de la longueur de l'encodage d'un nombre rationnel que:

$$\langle rs \rangle = \left\langle \frac{p}{q} \frac{t}{u} \right\rangle = \langle p \ t \rangle + \langle q \ u \rangle.$$

et donc, en utilisant i., que

$$\begin{aligned} \langle rs \rangle &\leq \langle p \rangle + \langle t \rangle + \langle q \rangle + \langle u \rangle \\ &= \langle p \rangle + \langle q \rangle + \langle t \rangle + \langle u \rangle \\ &= \langle r \rangle + \langle s \rangle. \end{aligned}$$

Nous obtenons ainsi l'inégalité (6).

b. Soit $D \in \mathbb{Z}^{n \times n}$. Par le point c. de la proposition 3.1, nous avons les équivalences suivantes:

$$\begin{aligned} |\det D| \leq 2^{\langle D \rangle - n^2} - 1 &\Leftrightarrow |\det D| + 1 \leq 2^{\langle D \rangle - n^2} \\ &\Leftrightarrow \log_2(|\det D| + 1) \leq \lceil \log_2(|\det D| + 1) \rceil \leq \langle D \rangle - n^2 \\ &\Leftrightarrow \langle \det D \rangle - 1 \leq \langle D \rangle - n^2 \end{aligned}$$

(par définition de la longueur de l'encodage d'un entier)

$$\Leftrightarrow \langle \det D \rangle \leq \langle D \rangle - n^2 + 1.$$

L'assertion (7) est donc montrée.

¹ p, q sont premiers entre eux.

² t, u sont premiers entre eux.

- c. Soit $D \in \mathbb{Q}^{n \times n}$. Comme les éléments de la matrice sont rationnels, ils peuvent se réécrire comme le rapport de deux entiers premiers entre eux.

Notons $(\frac{p_{ij}}{q_{ij}})_{i,j=1,\dots,n}$ les éléments de la matrice D .

$$\text{avec } \begin{matrix} p_{ij}, q_{ij} \in \mathbb{Z} & i, j = 1, \dots, n \\ q_{ij} \geq 1 \end{matrix}$$

Par le point c. de la proposition 3.1., nous avons les inégalités suivantes:

$$\begin{aligned} |\det D| &\leq 2^{\langle D \rangle - n^2} - 1 \\ &= 2^{\sum_{i,j=1}^n \langle p_{ij}/q_{ij} \rangle - n^2} - 1 \\ &= 2^{\sum_{i,j=1}^n (\langle p_{ij} \rangle + \langle q_{ij} \rangle) - n^2} - 1 \\ &\leq 2^{\sum_{i,j=1}^n \langle p_{ij} \rangle - n^2} - 1. \end{aligned} \quad (10)$$

Considérons ensuite l'entier $Q = \prod_{i,j=1}^n q_{ij}$.

Le déterminant de D peut s'écrire alors $\det D = (Q \det D) / Q$. (11)

Remarquons que $Q \det D \in \mathbb{Z}$.

En effet le déterminant de la matrice D est donné par l'expression suivante

$$\det D = \sum_{P \in S_n} \operatorname{sgn}(P) \frac{p_{1,P(1)}}{q_{1,P(1)}} \frac{p_{2,P(2)}}{q_{2,P(2)}} \dots \frac{p_{n,P(n)}}{q_{n,P(n)}}$$

où S_n est l'ensemble des permutations des entiers entre 1 et n et $\operatorname{sgn}(P)$ est la signature de la permutation $P \in S_n$.

En multipliant le déterminant de D par Q et en appliquant la distributivité, nous obtenons que $Q \det D$ est égal à une somme de produits d'entiers.

Il suit que $Q \det D \in \mathbb{Z}$.

L'obtention de ces résultats nous permet d'écrire les relations suivantes

$$\langle \det D \rangle = \langle Q \det D / Q \rangle \quad \text{par (11)}$$

$$= \langle Q \det D \rangle + \langle Q \rangle$$

$$= 1 + \lceil \log_2(|Q \det D| + 1) \rceil + \langle Q \rangle$$

$$\leq 1 + \lceil \log_2(Q (2^{\sum_{i,j=1}^n \langle p_{ij} \rangle - n^2} - 1) + 1) \rceil + \langle Q \rangle \quad \text{par (10)}$$

$$= 1 + \lceil \log_2(Q 2^{\sum_{i,j=1}^n \langle p_{ij} \rangle - n^2} - Q + 1) \rceil + \langle Q \rangle$$

$$\leq 1 + \lceil \log_2(Q 2^{\sum_{i,j=1}^n \langle p_{ij} \rangle - n^2}) \rceil + \langle Q \rangle \quad \text{car } Q \geq 1$$

$$\leq 1 + \lceil \log_2 Q + \log_2 2^{\sum_{i,j=1}^n \langle p_{ij} \rangle - n^2} \rceil + \langle Q \rangle$$

$$\leq 1 + \lceil \log_2 Q \rceil + \sum_{i,j=1}^n \langle p_{ij} \rangle - n^2 + \langle Q \rangle$$

$$\leq 1 + \lceil \log_2(Q + 1) \rceil + \sum_{i,j=1}^n \langle p_{ij} \rangle - n^2 + \langle Q \rangle$$

$$(Q \leq Q + 1)$$

$$= 2 \langle Q \rangle + \sum_{i,j=1}^n \langle p_{ij} \rangle - n^2$$

$$= 2 \left\langle \prod_{i,j=1}^n q_{ij} \right\rangle + \sum_{i,j=1}^n \langle p_{ij} \rangle - n^2$$

$$\leq 2 \sum_{i,j=1}^n \langle q_{ij} \rangle + \sum_{i,j=1}^n \langle p_{ij} \rangle - n^2 \quad \text{par (6)}$$

$$\leq 2 \sum_{i,j=1}^n \langle q_{ij} \rangle + 2 \sum_{i,j=1}^n \langle p_{ij} \rangle - n^2$$

$$= 2 \langle D \rangle - n^2$$

Ainsi $\langle \det D \rangle \leq 2 \langle D \rangle - n^2$, ce qui termine la preuve de la proposition 3.2. ■

Pour le problème de la programmation linéaire que nous avons défini dans l'introduction,

$$\left[\begin{array}{ll} \text{Minimiser} & c^T x \\ \text{sous} & A x = b \\ & x \geq 0 \end{array} \right.$$

nous supposons que toutes les données sont entières; seuls des entiers d'une longueur limitée peuvent être calculés exactement. Nous écrivons $L \equiv \langle A, b, c \rangle$ pour noter la longueur de l'encodage du problème. Il découle des propositions précédentes que 2^L est une borne supérieure de toute opération (addition, multiplication, calcul de déterminant...) effectuée sur les données du problème, sans répétition.

4. La complexité polynomiale d'un algorithme.

4.1. Définitions

- a. *La fonction de complexité* ou la fonction de temps d'exécution d'un algorithme est une fonction $f: \mathbb{N} \rightarrow \mathbb{N}$ exprimant le temps maximal $f(n)$ nécessaire pour résoudre un problème dont la longueur de l'encodage est au plus $n \in \mathbb{N}$. Le "temps" signifie le nombre de pas pour atteindre la solution du problème.
- b. *Un algorithme polynomial* est un algorithme dont la fonction de complexité satisfait $f(n) \leq p(n) \forall n \in \mathbb{N}$, pour un certain polynôme p . La complexité d'un tel algorithme croît *polynomialement* avec la taille de l'entrée.

Nous pouvons encore définir un algorithme polynomial comme un algorithme pour lequel la complexité asymptotique est un polynôme ou même, est bornée par un polynôme. Des algorithmes de complexité $O(n)$, $O(n^3)$ sont polynomiaux.

Pour mieux comprendre la signification d'un algorithme borné polynomialement, considérons les algorithmes restant, ceux qui violent toutes les bornes polynomiales: les algorithmes exponentiels.

Nous définissons les algorithmes exponentiels comme des algorithmes dont la complexité asymptotique est une fonction exponentielle ou est bornée par une fonction exponentielle.

Des exemples d'ordre de grandeur exponentiel sont $k^n (k > 1)$, $n^{\log n}$, $n!$

Un des algorithmes les plus connus ayant une complexité exponentielle est l'algorithme du simplexe. Ceci sera illustré dans le paragraphe 5.

Pour terminer, nous donnons quelques caractéristiques de l'algorithme polynomial qui rendent celui-ci plus performant que l'algorithme exponentiel.

4.2. Quelques caractéristiques ...

Le tableau de la *figure a* montre qu'une augmentation de la taille des données d'un problème, soit n , rend tout algorithme polynomial plus efficace qu'un algorithme exponentiel.

fonction	valeurs approximatives		
n	10	100	1000
$n \log n$	33	664	9966
n^3	1000	1000000	10^9
$10^6 n^8$	10^{14}	10^{22}	10^{30}
2^n	1024	$1,27 \cdot 10^{30}$	$1,05 \cdot 10^{301}$
$n^{\log n}$	2099	$1,93 \cdot 10^{13}$	$7,89 \cdot 10^{29}$
$n!$	3628800	10^{158}	$4 \cdot 10^{2567}$

figure a.

Le tableau de la *figure b* met en évidence qu'une découverte technologique est plus profitable à un algorithme polynomial qu'à un algorithme exponentiel. Par exemple, si la vitesse des ordinateurs décuple, la taille d'un problème résolu par un algorithme polynomial en un jour sera multipliée par une constante entre 1 et 10. Tandis que pour un algorithme exponentiel, nous n'aurons qu'une augmentation additive dans la taille du problème.

fonction	taille du problème résolu en un jour	taille du problème résolu en un jour par un ordinateur dix fois plus rapide
n	10^{12}	10^{13}
$n \log n$	$0,948 \cdot 10^{11}$	$0,87 \cdot 10^{12}$
n^2	10^6	$3,16 \cdot 10^6$
n^3	10^4	$2,15 \cdot 10^4$

2^n	40	43
10^n	12	13
$n^{\log n}$	79	95
$n!$	14	15

figure b.

5. L'algorithme du simplexe: complexité exponentielle

L'idée générale de l'algorithme du simplexe appliqué à un problème de programmation linéaire est la suivante:

L'algorithme démarre en un sommet du polyèdre défini par les contraintes du problème et vérifie s'il existe une arête contenant ce sommet et le long de laquelle la valeur de la fonction objectif peut diminuer. Si ce n'est pas le cas, le sommet courant est une solution optimale du problème. S'il y a de telles arêtes, nous suivons la direction de l'une d'elles. Soit nous rencontrons un nouveau sommet, auquel cas nous répétons la procédure ou alors l'arête est dite "infinie", auquel cas le problème est non borné.

Dans le cadre des problèmes bornés, il est possible de justifier la complexité exponentielle de l'algorithme du simplexe par le choix particulier d'une arête donnant lieu à une direction de descente.

A cet effet, nous présentons l'exemple de Klee et Minty. Nous considérons le polytope

défini par les inégalités $0 \leq x_j \leq 1 \quad j = 1, 2, 3$.

Il s'agit d'un cube (2 x 3 faces, 2^3 sommets) *figure c*.

figure c.

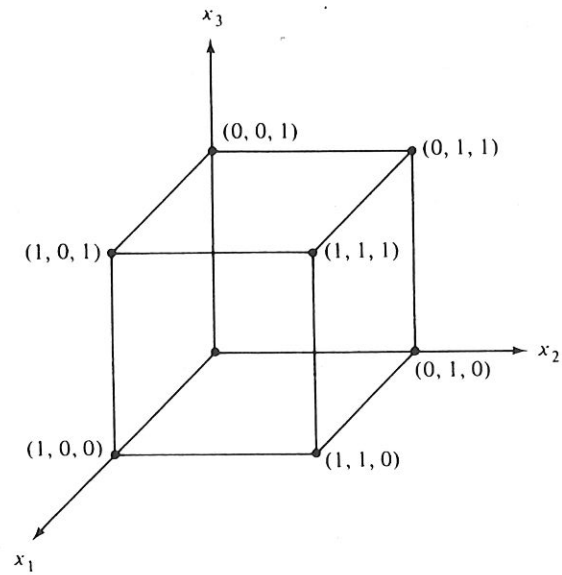
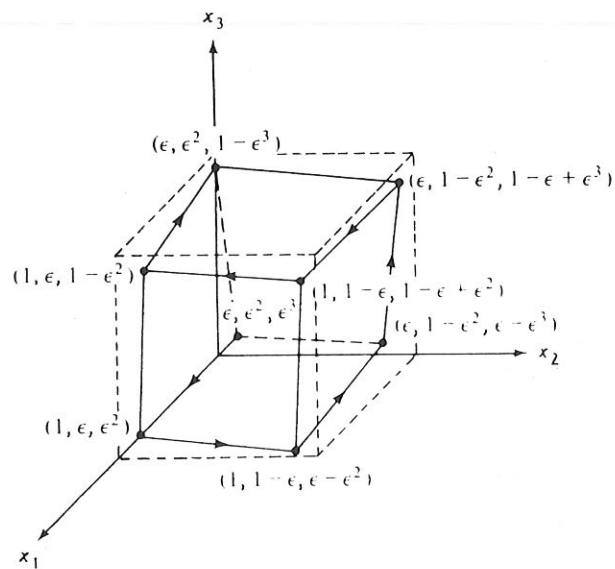


figure d.



Nous supposons que nous orientons le polytope de telle façon que le coût diminue lorsque l'on se dirige vers le haut. Il s'agira donc de déterminer une suite exponentielle de sommets adjacents, chacun étant plus élevé que le précédent. Par une "perturbation" du cube, nous construisons le polytope défini par les inégalités suivantes :

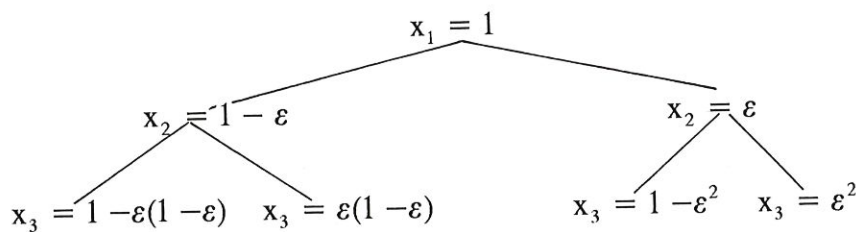
$$\begin{aligned} 1 &\geq x_1 \geq \varepsilon \\ 1 - \varepsilon x_{j-1} &\geq x_j \geq \varepsilon x_{j-1} \end{aligned} \quad j = 2, 3 \quad (12)$$

où $0 < \varepsilon < 1/2$.

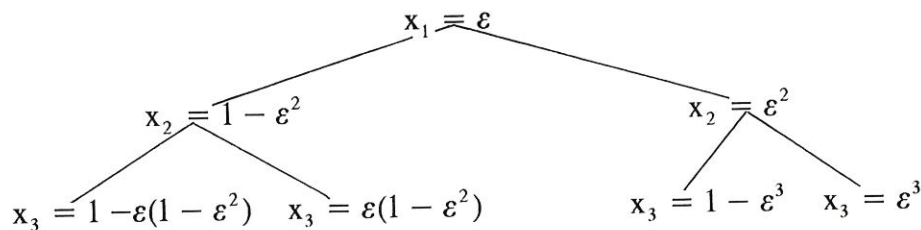
Le cube est donc la limite de ce polytope lorsque nous faisons tendre ε vers 0.

Vérifions que ces inégalités donnent lieu aux sommets de la figure d..

Par (12), les composantes des sommets peuvent être représentées par la structure "d'arbre" suivante:



ou



dès lors, les sommets sont

$$\begin{aligned} &(1, 1 - \varepsilon, 1 - \varepsilon + \varepsilon^2), (1, 1 - \varepsilon, \varepsilon - \varepsilon^2) \\ &(1, \varepsilon, 1 - \varepsilon^2), (1, \varepsilon, \varepsilon^2), (\varepsilon, 1 - \varepsilon^2, 1 - \varepsilon + \varepsilon^3) \\ &(\varepsilon, 1 - \varepsilon^2, \varepsilon - \varepsilon^3), (\varepsilon, \varepsilon^2, 1 - \varepsilon^3), (\varepsilon, \varepsilon^2, \varepsilon^3) \end{aligned}$$

En vue de montrer la complexité exponentielle de l'algorithme du simplexe, nous établissons l'existence d'une suite exponentielle de sommets adjacents.

Il est important de rappeler que l'ensemble des sommets joints par une arête correspond à l'ensemble des solutions de base admissibles adjacentes, non dégénérées³.

Il suffit donc de montrer l'existence d'une suite exponentielle de solutions de base admissibles adjacentes, non dégénérées.

³Théorème 2.10 cfr[10].

Suivant nos conventions, le problème consiste à maximiser la dernière composante, x_3 , sous les contraintes (12). La forme standard de ce problème est donnée, en considérant les variables d'écart r et s , par:

$$(PS) \equiv \left[\begin{array}{ll} \text{Minimiser} & -x_3 \\ \text{sous} & \begin{array}{l} x_1 - r_1 = \varepsilon \\ x_1 + s_1 = 1 \\ x_j - \varepsilon x_{j-1} - r_j = 0 \\ x_j + \varepsilon x_{j-1} + s_j = 1 \quad j = 2, 3 \\ x_j, r_j, s_j \geq 0 \quad j = 2, 3 \end{array} \end{array} \right.$$

Nous disposons d'un système de six équations à neuf inconnues.

Quelles sont les solutions de base réalisables?

La proposition suivante répond à la question

Proposition 5.1.

Toute solution de base réalisable est de la forme $(x_1, x_2, x_3, r_1, r_2, r_3, s_1, s_2, s_3)$, avec

- $x_1, x_2, x_3 > 0$
 - *exactement un des r_j, s_j est > 0 et l'autre nul et cela pour $j = 1, 2, 3$*
- (13)

Démonstration

Soit $(x_1, x_2, x_3, r_1, r_2, r_3, s_1, s_2, s_3)$ une solution de base réalisable.

a. Vérifions que x_1, x_2, x_3 sont strictement positifs:

Une solution de base réalisable vérifie les contraintes du problème (PS); d'où par la première contrainte, nous avons :

$$x_1 - r_1 = \varepsilon.$$

Or, la contrainte de signe nous dit que $r_i \geq 0$, dès lors,

$$x_1 \geq \varepsilon.$$

Par la troisième contrainte, nous avons:

$$x_{j+1} - \varepsilon x_j - r_{j+1} = 0 \quad \text{pour } j = 1, 2$$

c'est-à-dire

$$x_{j+1} - \varepsilon x_j = r_{j+1}.$$

Or $r_{j+1} \geq 0$, donc

$$x_{j+1} \geq \varepsilon x_j.$$

Il suit que $x_j \geq \varepsilon^j > 0$ pour $j = 1, 2$ ou 3 .

Les variables x_j sont donc des variables en base.

b. Supposons qu'il existe $j \in \{1, 2, 3\}$ tel que $r_j = s_j = 0$.

Envisageons deux possibilités:

1. $j = 1$

Comme $r_1 = s_1 = 0$, les deux premières contraintes nous apprennent que $x_1 = \varepsilon$ et $x_1 = 1$.

Ceci est impossible puisque $0 < \varepsilon < \frac{1}{2}$.

2. $j > 1$

Comme $r_j = s_j = 0$ pour $j = 2$ ou 3 , nous obtenons par les contraintes 3 et 4

$$x_j = \varepsilon x_{j-1} \quad \text{et} \quad x_j + \varepsilon x_{j-1} = 1.$$

Ainsi, nous avons

$$2 \varepsilon x_{j-1} = 1.$$

Or les contraintes 2 et 4 nous apprennent que :

$$x_1 \leq 1 \quad \text{et} \quad x_j \leq 1 \quad j = 2, 3$$

d'où

$$x_{j-1} \leq 1 \quad j = 2, 3.$$

Nous obtenons donc

$$x_{j-1} = \frac{1}{2\varepsilon} \leq 1.$$

Mais ceci est impossible puisque

$$\varepsilon < \frac{1}{2}.$$

Comme chaque cas nous mène à une contradiction, nous pouvons en conclure que l'hypothèse de départ est fausse et donc que, pour tout $j = 1, 2, 3$, r_j et s_j ne peuvent s'annuler en même temps.

c. En conclusion, les composantes x_1, x_2, x_3 strictement positives sont des variables en base.

De plus chaque couple (r_j, s_j) ($j = 1, 2, 3$) ne peut s'annuler et donc exactement un des r_j ou s_j est une variable de base, pour $j = 1, 2, 3$. ■

La proposition suivante nous permet d'affirmer que seules les solutions de base admissibles sont de la forme (13).

Proposition 5.2.

Tout vecteur admissible de la forme $(x_1, x_2, x_3, r_1, r_2, r_3, s_1, s_2, s_3)$ tel que

- $x_1, x_2, x_3 > 0$
- *exactement un des r_j, s_j est > 0 et l'autre nul et cela pour $j = 1, 2, 3$.*

est univoquement déterminé et est solution de base admissible.

Démonstration

Tout vecteur admissible vérifiant les hypothèses est solution du système

$$B_s^{-1} \begin{bmatrix} \varepsilon \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

où B_s est une sous-matrice carrée de la matrice A

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -\varepsilon & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -\varepsilon & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \varepsilon & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & \varepsilon & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

B_s doit comprendre les trois premières colonnes et exactement une colonne correspondant à s_j ou r_j pour $j = 1, 2, 3$.

On note S le sous-ensemble de $\{1, 2, 3\}$ qui correspond aux indices des colonnes de B correspondant à r .

Il est facile de montrer que chacune de ces sous-matrices est non singulière et correspond donc à une base admissible. La non singularité de ces matrices nous permet d'affirmer l'unicité de la solution du système correspondant. ■

Nous écrivons une solution de base admissible, x^s , où S est le sous-ensemble de $\{1, 2, 3\}$ qui correspond aux composantes non nulles de r dans x^s . La valeur de x_j dans x^s est notée x_j^s .

Lemme 5.3.

- a. Soit $j \in \{1, 2, 3\}$,
si $j \in S$ et $j \notin S'$ alors $x_j^S > x_j^{S'}$.
- b. Soit $j \in \{2, 3\}$,
si $S' = S - \{j\}$ alors $x_j^{S'} = 1 - x_j^S$.

Démonstration

a. Envisageons les deux cas suivants :

1. $j = 1$

Comme $1 \in S$, nous avons par définition de S , $r_1 \neq 0$ et donc $s_1 = 0$.
Par la deuxième contrainte, nous avons:

$$x_1^S = 1.$$

Or $1 \notin S'$, donc $r_1 = 0$.

Il suit de la première contrainte que

$$x_1^{S'} = \varepsilon$$

et comme $0 < \varepsilon < \frac{1}{2}$, nous avons bien

$$x_1^S > x_1^{S'}.$$

2. $j > 1$

i. Montrons d'abord que $x_j^S > \frac{1}{2}$. Comme $j \in S$, nous avons par définition de S ,
 $r_j \neq 0$ et donc $s_j = 0$. Par la quatrième contrainte, nous obtenons

$$x_j^S = 1 - \varepsilon x_{j-1}^S.$$

Or $x_{j-1}^S \leq 1$ et $\varepsilon < \frac{1}{2}$. Dès lors, nous en déduisons que

$$x_j^S = 1 - \varepsilon x_{j-1}^S > \frac{1}{2}.$$

- ii. Montrons ensuite que $x_j^{s'} < \frac{1}{2}$. Comme $j \notin S'$, nous avons $r_j = 0$. Nous obtenons de cette façon, en utilisant la troisième contrainte,

$$x_j^{s'} = \varepsilon x_{j-1}^{s'}.$$

Or $x_{j-1}^{s'} \leq 1$ et $\varepsilon < \frac{1}{2}$. Nous pouvons en déduire que

$$x_j^{s'} < \frac{1}{2}.$$

En conclusion, les points i et ii nous permettent d'affirmer que

$$x_j^{s'} < x_j^s \quad \text{pour } j = 2 \text{ ou } 3.$$

La première partie du lemme est ainsi démontrée.

- b. Faisons la preuve de l'implication pour $j = 3$, le cas où $j = 2$ est analogue.

Commençons par montrer que $x_1^{s'} = x_1^s$ et $x_2^{s'} = x_2^s$:

i. $x_1^{s'} = x_1^s$

Puisque $S = S' \cup \{3\}$ par hypothèse, $1 \in S'$ implique que $1 \in S$.

Dès lors, par définition de S , $r_1 > 0$ et $s_1 = 0$.

La deuxième contrainte nous conduit à

$$x_1^s = 1 = x_1^{s'}.$$

D'autre part, si $1 \notin S'$, nous avons $1 \notin S$. Dès lors, $r_1 = 0$ et $s_1 > 0$.

Nous concluons, par la première contrainte que

$$x_1^{s'} = \varepsilon = x_1^s.$$

ii. $x_2^{s'} = x_2^s$

L'hypothèse $S = S' \cup \{3\}$ implique que si $2 \in S'$, alors $2 \in S$.

Nous avons alors $r_2 > 0$ et $s_2 = 0$.

En utilisant le point i. et la quatrième contrainte, nous obtenons

$$x_2^s = 1 - \varepsilon x_1^s = 1 - \varepsilon x_1^{s'} = x_2^{s'}.$$

Par ailleurs, si $2 \notin S'$ alors $2 \notin S$ et donc $r_2 = 0$, $s_2 > 0$.

Le point i. et la troisième contrainte nous conduisent alors aux relations :

$$x_2^s = \varepsilon x_1^s = \varepsilon x_1^{s'} = x_2^{s'}.$$

Finalement, nous obtenons $x_3^{s'} = 1 - x_3^s$ en considérant :

$$\begin{aligned} x_3^{s'} &= \varepsilon x_2^{s'} \text{ car } 3 \notin S' \text{ (} r_3 = 0 \text{)} \\ &= 1 - (1 - \varepsilon x_2^{s'}) \\ &= 1 - (1 - \varepsilon x_2^s) \text{ par ii.} \\ &= 1 - x_3^s \text{ car } 3 \in S \text{ (} s_3 = 0 \text{)}. \quad \blacksquare \end{aligned}$$

Proposition 5.4.

Considérons les sous-ensembles de $\{1, 2, 3\}$, soient S_1, S_2, \dots, S_8 , énumérés de telle façon que $x_3^{s_1} \leq \dots \leq x_3^{s_8}$.

Nous pouvons montrer que les inégalités sont strictes et que les solutions de base admissibles x^{s_j} et $x^{s_{j+1}}$ sont adjacentes, pour $j = 1, \dots, 7$.

Démonstration

Nous procédons par différentes étapes. Montrons dans un premier temps que:

1. $x_1^{s_1} < x_1^{s_2}$ où S_1 et S_2 sont les sous-ensembles de $\{1\}$.

Soient $S_1 = \{\emptyset\}$ et $S_2 = \{1\}$. Nous avons alors deux solutions de base admissibles (x_1, r_1, s_1) qui sont : $(\varepsilon, 0, 1 - \varepsilon)$ et $(1, 1 - \varepsilon, 0)$.

Dès lors, nous avons effectivement $x_1^{s_1} < x_1^{s_2}$ puisque $0 < \varepsilon < \frac{1}{2}$.

De plus, de toute évidence, x^{s_1} et x^{s_2} sont adjacents.

Montrons ensuite:

2. $x_2^{s_1} < x_2^{s_2} < x_2^{s_3} < x_2^{s_4}$ où S_1, S_2, S_3, S_4 sont les sous-ensembles de $\{1, 2\}$.

- i. $S_1 = \{\emptyset\}$ et $S_2 = \{1\}$ sont aussi des sous-ensembles de $\{1, 2\}$. Comme 2 n'appartient ni à S_1 , ni à S_2 , nous avons $r_2 = 0$ et donc, par la troisième contrainte,

$$x_2^{s_j} = \varepsilon x_1^{s_j} \text{ pour } j = 1, 2.$$

Par ailleurs, nous savons que $x_1^{s_1} < x_1^{s_2}$ (étape 1) et en multipliant les deux membres de l'inégalité par ε , il vient alors que

$$\varepsilon x_1^{s_1} < \varepsilon x_1^{s_2} \text{ c'est - à - dire } x_2^{s_1} < x_2^{s_2}. \quad (13)$$

ii. Considérons à présent, les sous-ensembles restant de $\{1, 2\}$.

Soient S'_1 et S'_2 tels que $S'_j = S_j \cup \{2\}$ pour $j = 1, 2$.

Comme 2 appartient à S'_j et n'appartient pas à S_j , nous obtenons, en vertu du lemme 5.3 que

$$x_2^{s'_j} > x_2^{s_2} \quad \text{pour } j = 1, 2. \quad (14)$$

De plus

$$S'_j = S_j \cup \{2\} \quad \text{pour } j = 1, 2.$$

D'où, à nouveau par le lemme 5.3, nous avons

$$x_2^{s'_j} = 1 - x_2^{s_j} \quad \text{pour } j = 1, 2. \quad (15)$$

Finalement, nous obtenons consécutivement, par (13) (14) (15)

$$x_2^{s_1} < x_2^{s_2} < x_2^{s'_2} < x_2^{s'_1}.$$

Par ailleurs, nous savons déjà que x^{s_1} et x^{s_2} sont adjacents et $S'_j = S_j \cup \{2\}$ pour $j = 1, 2$; donc $x^{s'_2}$ et $x^{s'_1}$ sont adjacents et x^{s_2} et $x^{s'_2}$ le sont aussi puisque S_2 et S'_2 diffèrent d'un seul indice.

Notons $S'_2 \equiv S_3$ et $S'_1 \equiv S_4$.

Montrons à présent que

3. $x_3^{s_1} < \dots < x_3^{s_8}$ où $S_1 \dots S_8$ sont des sous-ensembles de $\{1, 2, 3\}$

i. S_1, S_2, S_3, S_4 sont aussi des sous-ensembles de $\{1, 2, 3\}$.

Comme 3 n'appartient pas à S_1, \dots, S_4 , nous avons $r_3 = 0$ et donc

$$x_3^{s_j} = \varepsilon x_2^{s_j} \quad \text{pour } j = 1, 2, 3, 4.$$

Par l'étape 2, nous savons

$$x_2^{s_1} < x_2^{s_2} < x_2^{s_3} < x_2^{s_4}.$$

En multipliant cette série d'inégalités par ε nous obtenons

$$x_3^{s_1} < x_3^{s_2} < x_3^{s_3} < x_3^{s_4}. \quad (16)$$

ii. Considérons les sous-ensembles restant de $\{1, 2, 3\}$ et appelons-les S'_1, S'_2, S'_3 et S'_4 ,

Ils sont tels que $S'_j = S_j \cup \{3\}$ pour $j = 1, 2, 3, 4$.

Appliquons maintenant le lemme 5.3.:

Comme $3 \in S'_j$ et $3 \notin S_4$, nous en déduisons que $x_3^{s'_j} > x_3^{s_4}$ pour $j = 1 \dots 4$. (17)

De plus, nous avons $S'_j = S_j \cup \{3\}$. Il suit que

$$x_3^{s'_j} = 1 - x_3^{s_j}. \quad (18)$$

Nous obtenons, grâce à (16) (17) (18), la chaîne d'inégalités strictes:

$$x_3^{s'_1} < x_3^{s'_2} < x_3^{s'_3} < x_3^{s'_4} < x_3^{s'_5} < x_3^{s'_6} < x_3^{s'_7} < x_3^{s'_8}.$$

Finalement, nous savons que x^{s_j} et $x^{s_{j+1}}$ sont adjacents, pour $j = 1, 2, 3$.

De plus, $S'_j = S_j \cup \{3\}$ et $S'_{j+1} = S_{j+1} \cup \{3\}$, dès lors, $x^{s'_j}$ et $x^{s'_{j+1}}$ sont adjacents pour $j = 1, 2, 3$ et x^{s_4} et x^{s_5} le sont aussi, puisque la base correspondante à x^{s_4} est obtenue à partir de celle correspondant à x^{s_5} en retirant la colonne indiquée par s_3 et en ajoutant celle indiquée par r_3 .

En posant $S'_4 = S_5$, $S'_5 = S_6$, $S'_6 = S_7$ et $S'_7 = S_8$, nous obtenons la thèse. ■

Théorème 5.5.

L'algorithme du simplexe appliqué au problème (PS) de six équations à neuf inconnues peut prendre, en choisissant un pivotage particulier (c'est-à-dire par le choix particulier de r_j ou s_j comme variables de base), $2^3 - 1$ itérations pour atteindre l'optimum.

Démonstration

L'objectif du problème est de maximiser la troisième composante x_3 . La proposition 5.4. nous montre l'existence d'une série de solutions admissibles adjacentes formant une chaîne exponentielle.

En outre, cette chaîne augmente la valeur de x_3 et donc diminue le coût. L'optimum est donc atteint en $2^3 - 1$ itérations, c'est-à-dire le nombre de sommets moins un. ■

Remarque:

En dimension d où $d > 1$, nous obtenons $2^d - 1$ itérations. Ce qui donne bien une complexité de type exponentiel.

Chapitre II

Deux méthodes de minimisation en programmation non linéaire

1. La méthode "barrière", [11]

1.1. Présentation de la méthode

Nous désirons résoudre le problème de programmation non linéaire suivant :

$$(P) \equiv \begin{cases} \text{Minimiser } f(z) \\ \text{sous } z \in S \end{cases}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction continûment différentiable et S une partie de \mathbb{R}^n .
Nous faisons les hypothèses suivantes sur (P) :

$$1. S \text{ est fermé, d'intérieur non vide et } S = \overset{\circ}{S} \neq \emptyset \quad (19)$$

$$2. \text{ Il existe } \hat{z} \in S \text{ tel que l'ensemble } Z = \{ z \mid f(z) \leq f(\hat{z}) \} \text{ est compact.} \quad (20)$$

L'idée sous-jacente de la méthode barrière est de résoudre le problème (P) en générant une suite de points z^i de \mathbb{R}^n , optimaux pour une suite de problèmes de minimisation sans contrainte de la forme :

$$(Pe)_i \equiv \begin{cases} \text{Minimiser } f(z) + p_i(z) & \text{où } i = 0, 1, 2, \dots \\ z \in \overset{\circ}{S} \end{cases}$$

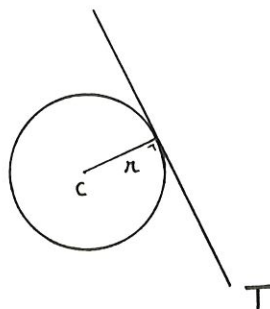
Ces sous-problèmes sont construits de telle manière que la suite (z^i) converge vers un élément z^* de S qui est optimal pour le problème initial (P).

¹Remarque : $\overset{\circ}{S}$ n'est pas toujours égal à S

Exemple : $S = B(c, \|r\|) \cup T$,

$$\overset{\circ}{S} = \overset{\circ}{B}(c, \|r\|),$$

$$\text{d'où } S \neq \overset{\circ}{S}$$



L'intérêt de cette méthode est qu'elle engendre une série de points intérieurs au domaine admissible.

Elle établit, à l'aide des fonctions p_i , une barrière le long de la frontière de S pour empêcher de quitter le domaine admissible, d'où vient son nom de méthode barrière.

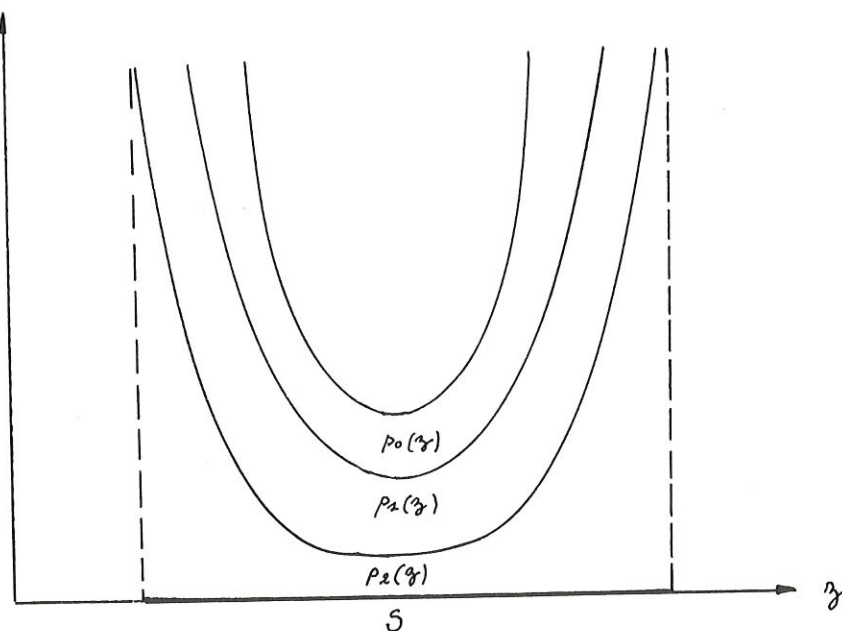
Une question se pose : comment construire les fonctions p_i ?

1.2. Les fonctions barrières

1.2.1. Définition

Une suite de fonctions continues $p_i : \overset{\circ}{S} \rightarrow \Re$ ($i = 0, 1, 2, \dots$) est appelée une suite de fonctions barrières pour l'ensemble S si les conditions suivantes sont satisfaites :

1. Pour tout point z intérieur à S , nous avons :
 $0 < p_{i+1}(z) < p_i(z)$ pour $i = 0, 1, 2, \dots$
 $p_i(z) \rightarrow 0$ lorsque $i \rightarrow \infty$
2. Pour toute suite (z^j) de points intérieurs à S , convergeant vers z^* , point frontière de S , nous avons:
 $p_i(z^j) \rightarrow \infty$ lorsque $j \rightarrow \infty$.



1.2.2. Exemples

Supposons que pour $j = 1, \dots, m$, $f^j : \mathbb{R}^n \rightarrow \mathbb{R}$ soit une fonction continue et considérons l'ensemble $S = \{z \mid f^j(z) \leq 0 \quad j = 1 \dots m\}$. Supposons ensuite que les hypothèses (19) et (20) soient satisfaites et que pour tout point z , intérieur à S , nous ayons $f^j(z) < 0 \quad j = 1, \dots, m$. Enfin, considérons une suite croissante de nombres strictement positifs que l'on note, $(\alpha_i) \quad i = 0, 1, 2 \dots$

Nous présentons deux exemples de fonction barrière pour l'ensemble S :

Exemple 1

$$p_i(z) = -\frac{1}{\alpha_i} \sum_{j=1}^m (1 / f^j(z))$$

Vérifions les différentes propriétés :

1. $0 < p_{i+1}(z) < p_i(z) \quad i = 0, 1, 2, \dots$

i. Par hypothèse, nous avons $\alpha_i > 0$ et $f^j(z) < 0$ pour $z \in \overset{\circ}{S}$ de sorte que

$$-\frac{1}{\alpha_i} \sum_{j=1}^m 1 / f^j(z) > 0 \quad \text{pour } z \in \overset{\circ}{S} \quad , \text{c'est-à-dire } p_i(z) > 0$$

pour $i = 0, 1, 2, \dots$

ii. Comme la suite (α_i) est croissante, nous avons $\alpha_{i+1} > \alpha_i$ ou encore

$$-\frac{1}{\alpha_{i+1}} > -\frac{1}{\alpha_i} \quad \text{pour tout } i = 0, 1, 2, \dots \text{ De plus } f^j(z) < 0 \text{ pour } z \in \overset{\circ}{S},$$

dès lors

$$-\frac{1}{\alpha_{i+1}} \sum_{j=1}^m 1 / f^j(z) < -\frac{1}{\alpha_i} \sum_{j=1}^m 1 / f^j(z) \quad \text{pour } i = 0, 1, 2, \dots$$

c'est-à-dire

$$p_{i+1}(z) < p_i(z) \quad \text{pour } i = 0, 1, 2, \dots$$

2. $p_i(z) \rightarrow 0$ lorsque $i \rightarrow +\infty$.

La suite (α_i) est croissante, donc $(\frac{1}{\alpha_i})$ converge vers 0 lorsque i tend vers ∞ .

Par conséquent $-\frac{1}{\alpha_i} \sum_{j=1}^m (1 / f^j(z)) \rightarrow 0$ lorsque $i \rightarrow \infty$.

3. Si $z^k \rightarrow z^*$ lorsque $k \rightarrow \infty$, où $z^* \in \text{front } S$, $p_i(z^k) \rightarrow \infty$ lorsque $k \rightarrow \infty$.

Comme f^j est une fonction continue pour $j = 1, 2, \dots, m$, nous avons

$$f^j(z^k) \rightarrow f^j(z^*) = 0 \text{ et donc } \frac{1}{f^j(z^k)} \rightarrow \infty.$$

Par conséquent, $-\frac{1}{\alpha_i} \sum_{j=1}^m 1 / f^j(z^k) \rightarrow \infty$.

Exemple 2

$$p_i(z) = -\frac{1}{\alpha_i} \sum_{j=1}^m \log[-f^j(z)]$$

Vérifions les différentes propriétés:

1. $0 < p_{i+1}(z) < p_i(z)$ pour tout $z \in \overset{0}{S}$ et $i = 0, 1, 2, \dots$

- La définition exacte de la fonction barrière est en réalité :

$$-\frac{1}{\alpha_i} \sum_{j=1}^m \log[-f^j(z) / M]$$

où $M = \max_j (-\min_{z \in S} f^j(z)) < \infty$.

Comme $\alpha_i > 0$ et $\frac{-f^j(z)}{M} < 1$, nous avons $-\frac{1}{\alpha_i} \sum_{j=1}^m \log[-f^j(z) / M]$ est > 0 ,

c'est-à-dire $p_i(z) > 0$ pour $i = 0, 1, 2, \dots$

- Comme la suite (α_i) est croissante, nous avons $\alpha_{i+1} > \alpha_i$ ou $\frac{-1}{\alpha_{i+1}} > \frac{-1}{\alpha_i}$.

De plus $\sum_{j=1}^m \log[-f^j(z) / M]$ est strictement négatif.

Dès lors $p_{i+1}(z) < p_i(z)$ où $z \in \overset{0}{S}$ et $i = 0, 1, 2, \dots$

Puisque le terme $\frac{1}{\alpha_i} M$ n'a pas d'effet sur le point optimal z_i , il peut être omis.

2. $p_i(z) \rightarrow 0$ quand $i \rightarrow \infty$ pour tout $z \in \overset{0}{S}$.

Il suffit d'utiliser la croissance de la suite (α_i) . Nous avons alors

$\alpha_i \rightarrow \infty$ quand $i \rightarrow \infty$ et donc $\frac{1}{\alpha_i}$ converge vers 0.

Ceci implique que $p_i(z) \rightarrow 0$ quand $i \rightarrow \infty$ pour tout $z \in \overset{0}{S}$.

3. Si $z^k \rightarrow z^*$, $z^* \in \text{front } S$, alors $p_i(z^k) \rightarrow \infty$ quand $k \rightarrow \infty$.

Il suit de la continuité de f^j que $f^j(z^k) \rightarrow f^j(z^*) = 0$ car $z^* \in \text{front } S$.

D'où $\log[-f^j(z^k)]$ converge vers $-\infty$ et donc $-\frac{1}{\alpha_i} \sum_{j=1}^m \log[-f^j(z^k)] \rightarrow +\infty$.

1.3. Convergence de la méthode.

L'existence d'un minimum pour chaque problème sans contrainte $(Pe)_i$ est assurée par le lemme suivant:

Lemme 1.3.1.

Considérons le problème pénalisé $(Pe)_i$. Supposons qu'il existe \hat{z} appartenant à S tel que l'ensemble $Z = \{z \mid f(z) \leq f(\hat{z}) + p_0(\hat{z})\}$ soit compact. Alors nous pouvons montrer que, pour $i = 0, 1, 2, \dots$, il existe un point intérieur z^i , $z^i \in \overset{0}{S}$, qui minimise $f(z) + p_i(z)$ pour tout $z \in \overset{0}{S}$.

Démonstration

Soit $\hat{z} \in \overset{\circ}{S}$ un point vérifiant les hypothèses du lemme.

Soit $S_i = \{z \in \overset{\circ}{S} \mid f(z) + p_i(z) \leq f(\hat{z}) + p_i(\hat{z})\}$ pour $i = 0, 1, 2, \dots$

1. Montrons l'existence d'un point z^i , appartenant à S_i , qui minimise $f(z) + p_i(z)$ sur $\overset{\circ}{S}$. Commençons par établir quelques propriétés de S_i :

a. $S_i \subset Z$

Soit $y \in S_i$. Montrons que $y \in Z$, c'est à dire que $f(y) \leq f(\hat{z}) + p_0(\hat{z})$.

Par la définition de S_i , nous avons: $f(y) + p_i(y) \leq f(\hat{z}) + p_i(\hat{z})$.

En utilisant les propriétés des fonctions barrières, $p_i > 0$ et $p_{i+1} < p_i$ pour $i = 0, 1, 2, \dots$, nous obtenons les inégalités suivantes :

$$f(y) < f(\hat{z}) + p_i(\hat{z}) \text{ et } f(y) < f(\hat{z}) + p_0(\hat{z}).$$

Par conséquent $y \in Z = \{z \mid f(z) \leq f(\hat{z}) + p_0(\hat{z})\}$.

b. S_i est fermé

Prenons une suite de points $(z^k)_{k=0}^\infty$ dans S_i convergente vers un élément z^* .

Montrons que ce dernier appartient à S_i , autrement dit, vérifions que

$$f(z^*) + p_i(z^*) \leq f(\hat{z}) + p_i(\hat{z}).$$

En effet, nous savons que z^k appartient à S_i pour tout k . Dès lors,

$$f(z^k) + p_i(z^k) \leq f(\hat{z}) + p_i(\hat{z}) \text{ pour tout } k.$$

Faisons tendre k vers l'infini et appliquons la propriété de continuité de f et de p_i . Nous obtenons de cette façon:

$$f(z^*) + p_i(z^*) \leq f(\hat{z}) + p_i(\hat{z}).$$

Autrement dit $z^* \in S_i$.

c. Conclusion

S_i est un ensemble fermé, inclus dans un ensemble compact, Z . Il existe donc un point qui minimise la fonction continue $f + p_i$ sur le compact S_i . Appelons z^i ce point.

2. z^i est optimal pour $(Pe)_i$.

Si nous considérons tous les points z de $\overset{\circ}{S}$, n'appartenant pas à S_i , nous avons la relation suivante:

$$f(z) + p_i(z) > f(\hat{z}) + p_i(\hat{z}).$$

Ces points ne minimisent pas la fonction $f + p_i$. Nous obtenons ainsi la thèse. ■

Théorème 1.3.2.

Soit z^i l'optimum pour le problème sans contrainte $(P_e)_i$, pour $i = 0, 1, 2, \dots$
Supposons que les hypothèses (19) et (20) soient satisfaites. Alors tout point
d'accumulation z^* de la suite $(z^i)_{i=0}^\infty$ est optimal pour le problème avec
contrainte (P) .

Démonstration

Soit b_i le minimum de $\{f(z) + p_i(z) \mid z \in \overset{\circ}{S}\}$ pour $i = 0, 1, 2, \dots$
Soit b le minimum de $\{f(z) \mid z \in S\}$.

Dans un premier temps, attachons-nous à démontrer que la suite (b_i) est convergente et que sa limite est le point b . Pour ce faire, procédons par étape:

1. La suite (b_i) , pour $i = 0, 1, 2, \dots$ est décroissante monotone:

En effet, par la propriété de la fonction barrière, nous avons:

$$f(z) + p_i(z) \geq f(z) + p_{i+1}(z) \quad \text{pour tout } z \in \overset{\circ}{S}, j = 0, 1, 2, \dots$$

C'est pourquoi,

$$\text{Min} \{f(z) + p_i(z) \mid z \in \overset{\circ}{S}\} \geq \text{Min} \{f(z) + p_{i+1}(z) \mid z \in \overset{\circ}{S}\}.$$

C'est-à-dire, $b_i \geq b_{i+1}$ pour tout $i = 0, 1, 2, \dots$

2. La suite (b_i) , pour $i = 0, 1, 2, \dots$ est bornée:

En effet, pour tout $i = 0, 1, 2, \dots$ nous avons $f(z) + p_i(z) \geq f(z) \quad \forall z \in S$.

$$\text{D'où } \text{Min} \{f(z) + p_i(z) \mid z \in S\} \geq \text{Min} \{f(z) \mid z \in S\}.$$

Comme $\overset{\circ}{S} \subseteq S$,

il suit que $\text{Min} \{f(z) + p_i(z) \mid z \in \overset{\circ}{S}\} > \text{Min} \{f(z) \mid z \in S\}$, c'est-à-dire
 $b_i > b \quad \forall i = 0, 1, 2, \dots$

Puisque la suite (b_i) est bornée et décroissante monotone, par le théorème de Bolzano Weierstrass, cette suite est convergente. Appelons b' sa limite.

Il est évident que $b' \geq b$.

Dans la troisième étape de la démonstration, nous montrons que $b' = b$.

3. $b' = b$

Supposons par l'absurde que $b' > b$.

i. Soit \hat{z} un point optimal pour le problème (P).

En utilisant la propriété de continuité de f et l'hypothèse (19), $\bar{\hat{S}} = S$, nous pouvons affirmer l'existence d'une boule B centrée en \hat{z} telle que:

$$B \cap \hat{S} \neq \emptyset \quad \text{et} \\ f(z) < \frac{b' + b}{2} = b' - \left(\frac{b' - b}{2}\right) \quad \forall z \in B$$

ii. Soit z un point de $B \cap \hat{S}$.

Par définition des fonctions barrières, $p_i(z) \rightarrow 0$ lorsque $i \rightarrow +\infty$.

Il existe donc un rang N à partir duquel $p_i(z)$ est inférieur à $\frac{b' - b}{4}$.

En rassemblant les résultats des points i) et ii) nous obtenons:

$$\forall i \geq N \quad b_i \leq f(z) + p_i(z) < b' - \frac{b' - b}{4}.$$

Nous déduisons que $b_i < b'$, ce qui est contraire au fait que b' soit la limite de la suite décroissante (b_i) . L'hypothèse, $b' > b$, est donc fausse. Nous concluons que $b' = b$.

Jusqu'à présent nous avons montré que la suite $(b_i)_{i=0}^\infty$ converge vers le point b .

Considérons maintenant, z^* , un point d'accumulation de la suite $(z^i)_{i=0}^\infty$.

Autrement dit, il existe une sous-suite de (z^i) convergente vers z^* . Appelons

(z^i) , $i \in K \subset \{0, 1, 2, \dots\}$, cette sous-suite.

Nous devons montrer que z^* est optimal pour le problème (P).

Supposons le contraire; nous savons alors que $f(z^*) > b$.

Il résulte que la suite $\{[f(z^i) - b] + p_i(z^i)\}$, $i \in K$ ne peut converger vers 0.

En effet si $[f(z^i) - b] + p_i(z^i) \rightarrow 0$, nous aurions

$p_i(z^i) \rightarrow -[f(z^*) - b] < 0$, ce qui est contraire au fait que $p_i(z^i)$ soit positif pour tout i , par définition de la fonction barrière.

Cela contredit le fait que $b_i - b$ égal à $b_i - b'$ converge vers 0 lorsque $i \rightarrow \infty$.

Nous obtenons ainsi que $f(z^*) = b$, c'est-à-dire que z^* est optimal pour le problème (P). ■

1.4. Modification de la méthode

Le théorème du paragraphe précédent établit la propriété essentielle de la méthode barrière, à savoir: Sous certaines hypothèses, elle construit une suite qui converge vers l'optimum du problème initial (P).

Ce résultat sous-entend que l'on peut résoudre le problème de minimisation sans contrainte, $(Pe)_i$, exactement et en un temps fini.

Cependant, les méthodes utilisées pour résoudre ces problèmes sans contrainte, par exemple la méthode de plus forte pente calculent seulement des points stationnaires et prennent généralement un nombre infini d'itérations pour calculer ces points stationnaires.

En conséquence, nous devons utiliser des procédures de troncature qui donneront une approximation du minimum de $f + p_i$.

Nous décrivons ci-après l'algorithme modifié de la méthode barrière.

Supposons, pour cela, que l'ensemble S satisfait les hypothèses (19) et (20) et posons

$p_i = \varepsilon_i p$ où $\varepsilon_i = \frac{1}{\alpha_i}$ pour $i = 0, 1, 2, \dots$ une suite de fonctions barrières pour l'ensemble S .

Supposons aussi l'existence d'un point z^0 dans $\overset{\circ}{S}$ tel que l'ensemble

$$\{z \mid f(z) \leq f(z^0) + \frac{1}{\alpha_0} p(z^0)\} \text{ soit compact.}$$

Enfin nous supposons que les fonctions f et $p(\cdot)$ sont deux fois continûment différentiables. L'algorithme modifié de la méthode barrière s'énonce comme suit:

Etape 0:

Sélection des nombres $\varepsilon > 0$, $\beta > 1$ et du point z^0 appartenant à $\overset{\circ}{S}$ tel que

$$\{z \mid f(z) \leq f(z^0) + \varepsilon p(z^0)\} \text{ soit compact.}$$

Etape 1:

Initialisation : $\varepsilon_0 = \varepsilon$

$$j = 0$$

$$i = 0.$$

A l'itération i ,

Etape 2:

$$j = 0, z^j = z^i.$$

Etape 3:

Calcul de la direction notée $h(z^j; \varepsilon_i)$ égale à $-\left[\nabla f(z^j) + \varepsilon_i \nabla p(z^j)\right]$

Etape 4:

Test d'arrêt:

$$SI \|h(z^j, \varepsilon_i)\| \leq \varepsilon_i$$

ALORS,

- on met à jour le paramètre $\varepsilon_{i+1} = \frac{\varepsilon_i}{\beta}$,
- z^j est la solution du problème sans contrainte, $z^{i+1} = z^j$,
- on ajoute 1 à l'indice: $i = i + 1$,
- on retourne à l'étape 2.

SINON, nous allons à l'étape 5.

Etape 5:

Nous effectuons une recherche linéaire pour obtenir $z^{j+1} = z^j + \lambda^j h(z^j; \varepsilon_i)$,
 $j = j + 1$.

Nous retournons à l'étape 3.

Dans le cas de contraintes d'inégalité, nous disposons d'un théorème nous assurant la convergence d'un tel algorithme.

Théorème 1.4.1

Si nous considérons le problème suivant:

$$\left[\begin{array}{l} \text{Minimiser } f(z) \\ \text{sous } g(z) \leq 0 \end{array} \right.$$

où $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ est continûment différentiable. Alors la suite générée par l'algorithme converge vers un point z^ admissible et vérifiant la condition nécessaire d'optimalité, c'est-à-dire, il existe des multiplicateurs μ^0, \dots, μ^m non tous nuls, tels que*

- $\sum_{i=0}^m \mu^i \nabla g^i(z^*) = 0$
- $\mu^i g^i(z^*) = 0 \quad \text{pour } i = 1, \dots, m.$

2. La méthode S-S-D [4]

L'idée générale de la méthode S-S-D est de suivre la direction de plus forte pente, après avoir effectué une transformation de coordonnées appelée "scaling".

Dans la littérature anglaise, cette méthode porte le nom: "The scaling steepest descent method". Nous l'appellerons méthode S-S-D.

2.1. La transformation "scaling"

Un "scaling" appliqué au problème

$$(P) \equiv \begin{cases} \text{Minimiser } f(x) \\ \text{sous} & Ax = b \\ & x \geq 0 \end{cases}$$

où $f: \mathcal{R}^n \rightarrow \mathcal{R}$ est continûment différentiable,

$b \in \mathcal{R}^m$, A est une matrice de dimension $m \times n$, de rang plein ($m < n$),
est une transformation de coordonnées définie de \mathcal{R}^n dans \mathcal{R}^n , qui, à chaque vecteur x fait correspondre un vecteur y dans l'espace transformé tel que $y = D^{-1}x$, où D est une matrice diagonale positive.

Que devient le problème (P) dans l'espace transformé?

Une transformation "scaling" appliquée à une fonction $f: \mathcal{R}^n \rightarrow \mathcal{R}$
 $x \rightarrow f(x)$

donne lieu à une fonction $\bar{f}: \mathcal{R}^n \rightarrow \mathcal{R}$
 $y \rightarrow \bar{f}(y)$

où $\bar{f}(y) = f(Dy)$.

L'équation $Ax = b$ devient après transformation $\bar{A}y = b$ où $\bar{A} = AD$.
Dès lors, le problème (P) devient, après application d'un "scaling" :

$$(\bar{P}) \equiv \begin{cases} \text{Minimiser } f(Dy) \\ \text{sous} & ADy = b \\ & y \geq 0 \end{cases}$$

Par ailleurs, si on considère un point intérieur à S , noté x° , un "scaling" autour de x° est la transformation

$$X_o^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$x \rightarrow X_o^{-1} x$$

$$\text{où } X_o = \text{diag} (x_1^o \dots x_n^o)$$

Ce changement de variables permet de transporter x^o sur le vecteur unité, noté e . Ce type de transformation nous apparaîtra très utile dans la suite des chapitres. Nous verrons qu'il permet de simplifier de lourdes expressions apparaissant dans des démonstrations, tout en laissant invariant le problème.

Remarque

Toutes les variables associées au problème transformé sont notées par une barre.

2.2. Méthode de "plus forte pente"

Nous nous proposons de résoudre le problème de programmation non linéaire suivant:

$$(P) \equiv \begin{cases} \text{Minimiser } f(x) \\ \text{sous } x \in \overset{\circ}{S} \end{cases}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^1 ,

$$\overset{\circ}{S} = \{ x \in \mathbb{R}^n \mid Ax = b, x > 0 \},$$

$$A \in \mathbb{R}^{m \times n} \text{ de rang plein } (m < n).$$

Soit $x^o \in \overset{\circ}{S}$.

La direction de plus forte pente est la direction qui minimise l'approximation linéaire de f autour de x^o dans une boule unité centrée en x^o .

Autrement dit, la direction de plus forte pente est solution du problème:

$$(P') \equiv \begin{cases} \text{Minimiser } \nabla f(x^o)^T d \\ \text{sous } \|d\| \leq \varepsilon \\ A d = 0 \end{cases}$$

$$\text{où } 0 < \varepsilon \leq 1.$$

Proposition 2.2.1.

La solution du problème (P') est un multiple de $h = -P_A \nabla f(x^0)$.

où P_A est la projection orthogonale sur le noyau de A ,

$$P_A = I - A^T (A A^T)^{-1} A$$

Démonstration

Le problème (P') peut se réécrire sous une forme équivalente

$$(P'') \equiv \begin{cases} \text{Minimiser } \nabla f(x^0)^T d \\ \text{sous} \quad \frac{1}{2} \|d\|^2 \leq \frac{\varepsilon}{2} \\ \quad \quad -A d = 0 \end{cases}$$

Appelons $w \in \mathbb{R}^m$ les multiplicateurs de Lagrange associés à la contrainte $A d = 0$ et

$v \in \mathbb{R}$, le multiplicateur de Lagrange associé à la contrainte $\frac{1}{2} \|d\|^2 \leq \frac{\varepsilon}{2}$.

Vérifions les conditions d'optimalité associées au problème (P''). En d'autres termes, vérifions s'il existe $w \in \mathbb{R}^m$ et $v \in \mathbb{R}$ tels que les équations suivantes soient satisfaites :

$$\begin{aligned} \nabla f(x^0) - A^T w + v d &= 0 \\ v (\|d\| - \varepsilon) &= 0 \\ v &\geq 0 \\ A d &= 0 \\ \|d\| &\leq \varepsilon \end{aligned}$$

Nous pouvons extraire w en considérant la première équation

$$A^T w = \nabla f(x^0) + v d,$$

qui est équivalente à

$$A A^T w = A \nabla f(x^0) + A v d = A \nabla f(x^0).$$

Nous savons, par hypothèse, que A est de rang plein. Dès lors $(A A^T)$ est inversible. Nous pouvons ainsi obtenir:

$$w = (A A^T)^{-1} A \nabla f(x^0).$$

Les conditions d'optimalité s'écrivent maintenant:

$$P_A \nabla f(x^o) + v d = 0$$

$$v (\|d\| - \varepsilon) = 0$$

$$v \geq 0$$

$$A d = 0$$

$$\|d\| \leq \varepsilon$$

Si nous choisissons $v = \frac{1}{\varepsilon} \|P_A \nabla f(x^o)\| \geq 0$, la direction $d^o = - \frac{P_A \nabla f(x^o)}{v}$ vérifie les conditions d'optimalité et est donc solution du problème (P'). ■

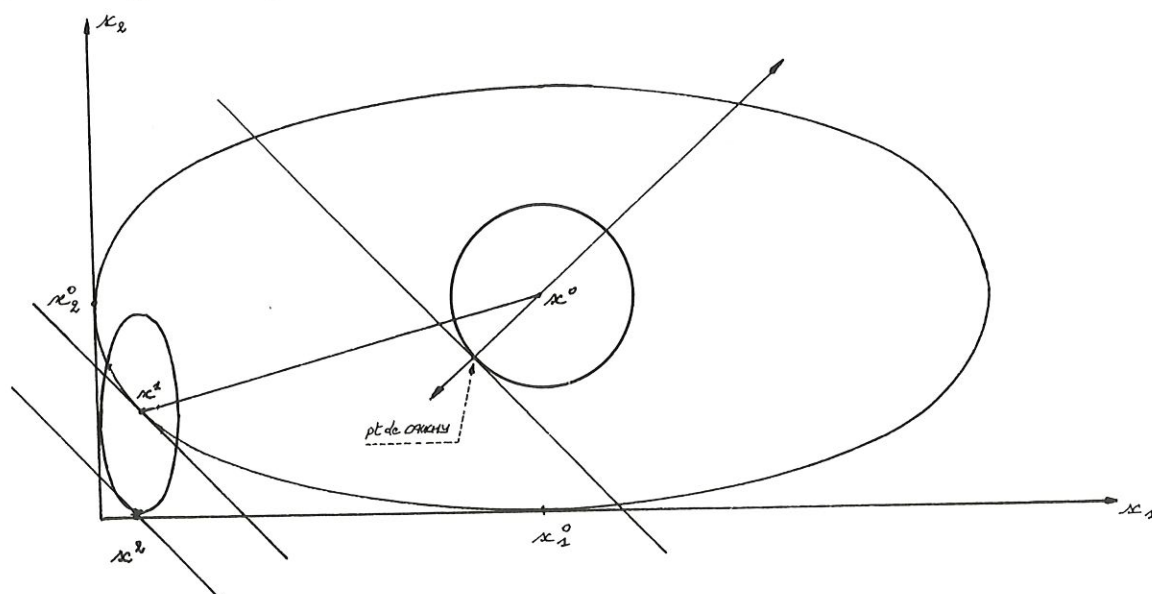
2.3. Utilité du scaling dans la méthode

Le calcul de la direction de plus forte pente correspond à un problème de minimisation dans une région de confiance. En effet, une fonction objectif simple, l'approximation linéaire de f , est minimisée dans une région simple, une boule, pour obtenir un modèle de la fonction autour du point.

Une boule est un choix évident pour une région de confiance car elle ne privilégie aucune direction. Mais la présence de contraintes de positivité motive la recherche d'une région qui rend compte plus précisément de la forme de la région admissible.

Si par exemple, la région admissible est l'orthant positif, \mathbb{R}_+^2 , un choix préférable est l'ellipsoïde simple, c'est-à-dire l'ellipsoïde dont les axes sont parallèles aux axes des coordonnées.

Représentation géométrique (à deux dimensions) :



En choisissant l'ellipsoïde comme région de confiance, la direction d^o sera alors solution du problème:

$$\left[\begin{array}{ll} \text{Minimiser} & \nabla f(x^o)^T d \\ \text{sous} & A d = 0 \\ & \sum_{i=1}^n d_i^2 / (x_i^o)^2 \leq \varepsilon \end{array} \right.$$

Ce problème peut se réécrire sous la forme équivalente

$$\left[\begin{array}{ll} \text{Minimiser} & \nabla f(x^o)^T d \\ \text{sous} & A d = 0 \\ & \|X_o^{-1} d\|^2 \leq \varepsilon \end{array} \right.$$

où $X_o = \text{diag}(x_1^o \dots x_n^o)$.

Effectuons, à présent, une transformation "scaling" autour du point x^o .

Le problème devient alors:

$$\left[\begin{array}{ll} \text{Minimiser} & \nabla \bar{f}(e)^T \bar{d} \\ \text{sous} & \bar{A} \bar{d} = 0 \\ & \|\bar{d}\|^2 \leq \varepsilon \end{array} \right.$$

où $\bar{d} = X_o^{-1} d$.

En effet, par la transformation $X_o = \text{diag}(x_1^o \dots x_n^o)$, nous obtenons:

$$\begin{aligned} \nabla f(x^o)^T d &= \nabla f(X_o e)^T d \\ &= \nabla f(X_o e)^T X_o X_o^{-1} d \\ &= \nabla \bar{f}(e)^T \bar{d} \\ \text{et } A d &= A X_o X_o^{-1} d = \bar{A} \bar{d}. \end{aligned}$$

Cette transformation permet de transporter l'ellipsoïde sur une boule centrée en x^o . Ainsi, ce dernier problème peut se résoudre comme le problème (P'). Nous obtenons

$$\begin{aligned} \bar{d}^o &= - \frac{\varepsilon}{\|P_{\bar{A}} \nabla \bar{f}(e)\|} P_{\bar{A}} \nabla \bar{f}(e) \\ &= - \frac{\varepsilon}{\|P_{\bar{A}} \nabla \bar{f}(e)\|} P_{A X_o} X_o \nabla f(x^o) \end{aligned}$$

2.4. L'algorithme S-S-D

Soient donnés $x^0 \in \overset{\circ}{S}$ et $f: \overset{\circ}{S} \rightarrow \Re$, une fonction continûment différentiable.

- Initialiser k à 0.
- Répéter les instructions suivantes jusqu'à ce que x^k converge vers la solution du problème.

— Transformation "scaling" autour de x^k :

$$\begin{aligned}\bar{A} &= A X_k \\ \bar{g} &= X_k \nabla f(x^k)\end{aligned}$$

— Calcul de la direction:

$$\begin{aligned}\bar{h} &= -P_{\bar{A}} \bar{g} \\ \text{où } P_{\bar{A}} &= I - \bar{A}^T (\bar{A} \bar{A}^T)^{-1} \bar{A} \\ &\text{est la matrice de projection sur le noyau de } \bar{A}.\end{aligned}$$

— Recherche linéaire le long de la direction \bar{h} :

Trouver $\bar{\lambda} \in \operatorname{argmin} \{ \bar{f}(e + \lambda \bar{h}) \mid \lambda \geq 0, e + \lambda \bar{h} \geq 0 \}$.

Nous obtenons $\bar{y} = e + \bar{\lambda} \bar{h}$, $\bar{y} > 0$.

— Retour dans l'espace d'origine par la transformation inverse du "scaling":

$$x^{k+1} = X_k \bar{y}$$

$$k = k + 1.$$

Chapitre III

La trajectoire centrale

1. Le problème de programmation linéaire

1.1. Présentation

Nous nous intéressons à la résolution du problème de programmation linéaire suivant :

$$(PL) \equiv \begin{cases} \text{Minimiser} & c^T x \\ \text{sous} & Ax = b \\ & x \geq 0 \end{cases}$$

où $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ et

A est une matrice de dimension (m,n) , de rang plein, où $0 < m < n$.

Nous supposons que l'ensemble admissible

$$S = \{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \}$$

est borné et son intérieur relatif

$$\overset{\circ}{S} = \{ x \in \mathbb{R}^n \mid Ax = b, x > 0 \}$$

est non vide.

Nous supposons également l'existence d'un point initial intérieur admissible, noté x° .

Remarque:

Nous définissons l'intérieur relatif d'un ensemble comme l'intérieur de cet ensemble relativement à la variété affine engendrée par cet ensemble.

Sous ces hypothèses, le problème a une solution optimale, notée \hat{x} , non nécessairement unique, et la valeur optimale du problème sera notée par

$$\hat{v} = c^T \hat{x}$$

1.2. Le concept de projection

Deux sous-espaces de \mathfrak{R}^n sont associés à la transformation linéaire dont la matrice est A :

- le noyau de A,

$$\text{Ker}(A) = \{ x \in \mathfrak{R}^n \mid Ax = 0 \}$$

- l'image de A^T , le complémentaire orthogonal du noyau de A,

$$\text{Im}(A^T) = \{ x \in \mathfrak{R}^n \mid x = A^T w, w \in \mathfrak{R}^m \}$$

Ainsi, tout vecteur d de \mathfrak{R}^n peut être décomposé de manière unique comme

$$d = dp + \hat{dp}$$

où $dp \in \text{Ker}(A)$ et $\hat{dp} \in \text{Im}(A^T)$,

dp est la projection de d sur le noyau de A et \hat{dp} , la projection de d sur le complémentaire orthogonal.

Nous avons $dp = P_A d$ et $\hat{dp} = \hat{P}_A d = (I - P_A) d$, où P_A est la matrice de projection sur le noyau de A.

Si A est une matrice de rang plein, la matrice de projection P_A est définie comme suit :

$$P_A = I - A^T (A A^T)^{-1} A.$$

Nous pouvons encore définir la projection de d sur le noyau de A comme le point dans $\text{Ker}(A)$ situé à une plus courte distance de d, c'est-à-dire,

$$dp = \text{argmin} \{ \|x - d\| \mid x \in \text{Ker}(A) \},$$

ou encore $dp = \text{argmin} \{ \|z\| \mid z = (d - A^T w), w \in \mathfrak{R}^m \};$

$$\|dp\| = \min \{ \|d - A^T w\| \mid w \in \mathfrak{R}^m \}.$$

Remarque :

Etant donné une matrice M, la matrice de projection sur le noyau de M est notée P_M . Si aucune confusion n'est possible, nous utilisons la notation simplifiée, P, pour la matrice de projection.

Par convention, nous notons r_p la projection d'un vecteur r sur le noyau de

$$M \quad (r_p = P_M r).$$

1.3. Le problème dual

Le problème dual associé au problème (PL) est

$$(D) \equiv \begin{cases} \text{Maximiser} & b^T w \\ \text{sous} & A^T w + z = c \\ & z \geq 0 \end{cases}$$

où $z \in \mathbb{R}^n$ est appelé variable duale d'écart et $w \in \mathbb{R}^m$.

Le problème (D) a une solution optimale, non nécessairement unique, notée (\hat{w}, \hat{z}) et nous avons $b^T \hat{w} = \hat{v}$.

Proposition 1.3.1

Soient $x \in S$ et (w, z) un couple admissible pour le problème (D).
Le saut de dualité est donné par

$$x^T z = c^T x - b^T w$$

Démonstration

Le saut de dualité est défini comme la différence entre le coût primal et le coût dual.
Autrement dit, il est égal à

$$c^T x - b^T w$$

En utilisant l'équation des contraintes du problème dual,

$$A^T w + z = c,$$

nous pouvons réécrire le saut de dualité comme suit

$$z^T x + w^T A x - b^T w.$$

Or $x \in S = \{ x \in \mathbb{R}^n / Ax = b, x \geq 0 \}$. Dès lors, nous obtenons

$$z^T x + w^T A x - b^T w = z^T x.$$

Ainsi le saut de dualité peut s'exprimer comme

$$x^T z = c^T x - b^T w. \quad \blacksquare$$

Proposition 1.3.2

Soit $z \in \mathbb{R}^n$

Les deux assertions suivantes sont équivalentes :

- i) z est admissible pour le problème dual (D)
- ii) $z \geq 0$ et $P_A z = P_A c$

Démonstration

Considérons un vecteur z de \mathbb{R}^n tel que $z \geq 0$.

Ce vecteur z est admissible pour (D), si et seulement si

$$c - z = A^T w \quad \text{où } w \in \mathbb{R}^m$$

$$\Leftrightarrow c - z \in \text{Im}(A^T)$$

Autrement dit $(c-z)$ est orthogonal au noyau de A , ou encore $P_A (c - z) = 0$.

Comme l'opérateur de projection est linéaire, nous obtenons la thèse

$$P_A c = P_A z. \quad \blacksquare$$

Etant donné x dans S , nous pouvons énoncer une forme équivalente du problème dual (D), (l'objectif diffère simplement d'une constante) :

$$(D') \equiv \begin{cases} \text{Minimiser} & x^T z \\ \text{sous} & P_A z = P_A c \\ & z \geq 0 \end{cases}$$

L'objectif est alors le saut de dualité et la valeur optimale est $x^T \hat{z} = c^T \hat{x} - \hat{v}$.

L'ensemble admissible de (D') est défini par

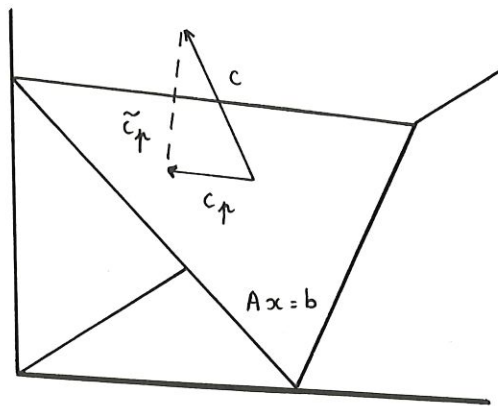
$$Z = \{ z \in \mathbb{R}^n \mid z \geq 0 \text{ et } P_A z = P_A c \}$$

et son intérieur relatif par

$$\overset{\circ}{Z} = \{ z \in \mathbb{R}^n \mid P_A z = P_A c \text{ et } z > 0 \}.$$

Revenons au problème primal (PL).

L'ensemble des solutions du problème (PL) ne change pas si nous remplaçons la fonction objectif par $c_p^T x$,



ou même, si nous considérons un vecteur z dual admissible, nous pouvons remplacer l'objectif par $z^T x$.
 Nous pouvons ainsi donner une forme équivalente au problème (PL) :

$$(PL') \equiv \begin{cases} \text{Minimiser} & z^T x \\ \text{sous} & Ax = b \\ & x \geq 0 \end{cases}$$

2. La fonction barrière logarithmique

2.1. Définition et propriétés immédiates

La fonction barrière logarithmique $p : \mathcal{R}_{++}^n \rightarrow \mathcal{R}$ est définie par

$$p(x) = - \sum_{i=1}^n \log x_i$$

Cette fonction pénalise les variables qui s'approchent de 0 et donc pénalise les points près du bord de la région admissible S .

Les dérivées de cette fonction sont :

$$\nabla p(x) = -x^{-1} \quad \nabla p(e) = -e \quad (21)$$

$$\nabla^2 p(x) = X^{-2} \quad \nabla^2 p(e) = I \quad (22)$$

où x^{-1} représente le vecteur dont la $i^{\text{ème}}$ composante est $\frac{1}{x_i}$,

$$X = \text{diag}(x_1, \dots, x_n),$$

$$e = [1 \dots 1]^T.$$

Comme la matrice hessienne de p , $\nabla^2 p(x)$, est définie positive dans $\overset{\circ}{S}$, la fonction barrière logarithmique est strictement convexe.

Nous pouvons remarquer la simplicité des dérivées de la fonction barrière évaluées au point e .

Cela nous motive à étudier ses propriétés autour du point e .
C'est l'objet du paragraphe suivant.

2.2. Variation de la fonction barrière autour du point e

Nous commençons par établir des résultats généraux sur la fonction logarithmique autour de 1.

Lemme 2.2.1

Soit $\lambda \in (-1, +1)$. Nous avons :

$$a. \log(1 + \lambda) \leq \lambda$$

$$b. \log(1 + \lambda) \geq \lambda - \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}$$

$$c. \frac{d}{d\lambda} \log(1 + \lambda) = \frac{1}{1 + \lambda} = 1 - \lambda + \frac{\lambda^2}{1 + \lambda}$$

Démonstration

a. Nous savons que la fonction logarithmique est concave.

Autrement dit, pour tout vecteur x et y de \mathfrak{R}^n , nous avons la relation suivante :

$$\log y \leq \log x + \frac{1}{x} (y - x). \quad (23)$$

Posons $y = 1 + \lambda$ et $x = 1$ et par substitution dans l'inégalité (23), nous obtenons

$$\log(1 + \lambda) \leq \lambda.$$

b. Comme la fonction logarithmique est différentiable, nous pouvons la développer en séries de Taylor.

Nous obtenons ainsi :

$$\log(1 + \lambda) = 1 - \frac{\lambda^2}{2} + \frac{\lambda^3}{3} - \frac{\lambda^4}{4} + \dots$$

ou encore

$$\log(1 + \lambda) = 1 - \frac{\lambda^2}{2} \left(1 - \frac{2}{3} \lambda + \frac{2}{4} \lambda^2 - \dots \right)$$

Nous pouvons minorer cette expression de sorte que

$$\log(1 + \lambda) \geq 1 - \frac{\lambda^2}{2} (1 + |\lambda| + |\lambda|^2 + \dots)$$

De plus, la série $1 + |\lambda| + |\lambda|^2 + \dots$ est une série géométrique de raison $|\lambda|$ dont la

somme vaut $\frac{1}{1 - |\lambda|}$. Finalement, l'inégalité devient

$$\log(1 + \lambda) \geq 1 - \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}.$$

c. La dernière propriété est immédiate. ■

Nous pouvons étendre facilement les propriétés du lemme 2.2.1 à la fonction barrière logarithmique.

Proposition 2.2.2

Soit $d \in \mathbb{R}^n$ un vecteur tel que $\|d\|_\infty < 1$.

Nous pouvons établir que :

$$a. \quad p(e + d) \geq \nabla p(e)^T d = -e^T d$$

$$b. \quad p(e + d) \leq \nabla p(e)^T d + \frac{\|d\|^2}{2} \frac{1}{1 - \|d\|_\infty}$$

$$= -e^T d + \frac{\|d\|^2}{2} \frac{1}{1 - \|d\|_\infty}$$

$$c. \quad \nabla p(e + d) = \nabla p(e) + d + o(d)$$

où $o(d) \in \mathbb{R}^n$ est tel que

$$\|o(d)\| \leq \frac{\|d\|^2}{1 - \|d\|_\infty}$$

Démonstration

a. Par l'hypothèse faite sur le vecteur d , nous avons $\|d\|_{\infty} < 1$.

Il suit que chaque composante du vecteur d appartient à l'intervalle $(-1, 1)$ pour tout $i = 1, \dots, n$.

Nous pouvons ainsi appliquer le lemme 2.2.1 et en particulier la première inégalité. Nous obtenons

$$\log(1 + d_i) \leq d_i \text{ pour tout } i.$$

En sommant les différentes inégalités, il vient que

$$\sum_{i=1}^n \log(1 + d_i) \leq \sum_{i=1}^n d_i,$$

ou encore, en multipliant chaque membre de l'inégalité par -1 ,

$$-\sum_{i=1}^n \log(1 + d_i) \geq -\sum_{i=1}^n d_i,$$

c'est-à-dire

$$p(e + d) \geq -e^T d.$$

b. En vertu du lemme 2.2.1, point b., nous savons que

$$\log(1 + d_i) \geq d_i - \frac{d_i^2}{2} \frac{1}{1 - |d_i|} \text{ pour } i = 1, \dots, n.$$

Cependant $|d_i| \leq \|d\|_{\infty}$, donc

$$\log(1 + d_i) \geq d_i - \frac{d_i^2}{2} \frac{1}{1 - \|d\|_{\infty}} \text{ pour tout } i.$$

Par sommation sur tous les indices $i = 1, \dots, n$, nous avons

$$\sum_{i=1}^n \log(1 + d_i) \geq \sum_{i=1}^n d_i - \frac{1}{2} \frac{1}{1 - \|d\|_{\infty}} \sum_{i=1}^n d_i^2.$$

Ceci est équivalent à

$$-\sum_{i=1}^n \log(1 + d_i) \leq -e^T d + \frac{\|d\|^2}{2} \frac{1}{1 - \|d\|_{\infty}},$$

c'est-à-dire

$$p(e + d) \leq -e^T d + \frac{\|d\|^2}{2} \frac{1}{1 - \|d\|_\infty}.$$

c. En utilisant la formule de la dérivée de la fonction barrière logarithmique (21),

$\nabla p(x) = -x^{-1}$, nous pouvons dire que

$$\nabla p(e + d) = - \begin{bmatrix} \frac{1}{1 + d_1} \\ \frac{1}{1 + d_2} \\ \vdots \\ \frac{1}{1 + d_n} \end{bmatrix}.$$

En vertu du point c. du lemme 2.2.1, nous savons que

$$\frac{1}{1 + d_i} = 1 - d_i + \frac{d_i^2}{1 + d_i}.$$

Dès lors, nous pouvons réécrire le gradient sous la forme

$$\nabla p(e + d) = -e + d - \begin{bmatrix} \frac{d_1^2}{1 + d_1} \\ \vdots \\ \frac{d_n^2}{1 + d_n} \end{bmatrix}$$

que l'on note encore :

$$\nabla p(e + d) = -e + d - \frac{d^2}{e + d}.$$

En outre, pour obtenir la thèse, il nous reste à montrer :

$$i. \quad \frac{-d^2}{e + d} = o(d)$$

En effet,

$$\lim_{d \rightarrow 0} \frac{-d^2}{e + d} = \lim_{d \rightarrow 0} \frac{-d}{e + d} = 0$$

$$\text{ii.} \quad \|o(d)\| \leq \frac{\|d\|^2}{1 - \|d\|_\infty}$$

En effet, comme $d_i \geq -\|d\|_\infty$, nous avons :

$$\frac{d_i^2}{1 + d_i} \leq \frac{d_i^2}{1 - \|d\|_\infty}$$

De cette inégalité, nous pouvons établir les relations suivantes :

$$\begin{aligned} \|o(d)\|^2 &= \sum_{i=1}^n \left(\frac{d_i^2}{1 + d_i} \right)^2 \\ &\leq \sum_{i=1}^n \frac{d_i^4}{(1 - \|d\|_\infty)^2} \\ &\leq \frac{\left(\sum_{i=1}^n d_i^2 \right)^2}{(1 - \|d\|_\infty)^2} \\ &= \frac{\|d\|^4}{(1 - \|d\|_\infty)^2} \end{aligned}$$

Nous vérifions ainsi la dernière assertion de la proposition. ■

2.3. Effet du " scaling " sur la fonction barrière logarithmique

Proposition 2.3.1

La transformation " scaling " laisse invariant toute variation de la fonction barrière logarithmique p.

Démonstration

Soit D une matrice diagonale positive, $D = \text{diag} (d_i)$.

Nous avons la décomposition suivante

$$\begin{aligned}
p(Dx) &= - \sum_{i=1}^n \log d_i x_i \\
&= - \sum_{i=1}^n \log x_i - \sum_{i=1}^n \log d_i \\
&= p(x) - \sum_{i=1}^n \log d_i
\end{aligned}$$

Par conséquent, pour tout x et y strictement positifs, nous obtenons
 $p(Dy) - p(Dx) = p(y) - p(x)$. ■

3. Les fonctions auxiliaires

Pour résoudre le problème de programmation linéaire (PL), nous avons supposé l'existence d'un point intérieur à la région admissible.
 Nous devons maintenant nous attacher à l'obtention d'un meilleur point qui respecte deux objectifs :

- Améliorer le coût, c'est-à-dire le diminuer.
- Rester le plus loin possible du bord du domaine admissible.

A cet effet, nous définissons différentes fonctions auxiliaires qui allient les deux objectifs :

- la fonction de pénalité intérieure,
- la fonction potentielle.

Chacune de ces fonctions a comme second terme la fonction barrière logarithmique permettant d'éviter le bord et est caractérisée par un paramètre qui rend compte d'une certaine manière de l'importance accordée à chacun des deux objectifs.

Nous réservons les deux paragraphes suivants aux définitions et propriétés des deux fonctions auxiliaires.

3.1. La fonction de pénalité intérieure

La fonction de pénalité intérieure est définie pour chaque valeur du paramètre $\alpha \in \mathbb{R}^+$ par

$$x \in \overset{\circ}{S} \rightarrow f_{\alpha}(x) = \alpha c^T x + p(x)$$

Ses dérivées sont

$$\nabla f_{\alpha}(x) = \alpha c - x^{-1} \quad \nabla f_{\alpha}(e) = \alpha c - e \quad (24)$$

$$\nabla^2 f_{\alpha}(x) = X^{-2} \quad \nabla^2 f_{\alpha}(e) = I \quad (25)$$

La fonction de pénalité intérieure diffère de la fonction barrière logarithmique p par le terme linéaire, $\alpha c^T x$.

En conséquence, comme la fonction barrière logarithmique, la fonction de pénalité intérieure est strictement convexe dans $\overset{\circ}{S}$.

De plus, les propriétés de variation autour du point e , énumérées dans la proposition 2.2.2 peuvent être étendues à la fonction de pénalité intérieure.

Proposition 3.1.1

Soit h un vecteur de \mathcal{R}^n tel que $\|h\| = 1$ et $\lambda \in [0, 1)$

Nous pouvons établir :

$$a. \quad f_{\alpha}(e + \lambda h) \geq f_{\alpha}(e) + \lambda \nabla f_{\alpha}(e)^T h$$

$$b. \quad f_{\alpha}(e + \lambda h) \leq f_{\alpha}(e) + \lambda \nabla f_{\alpha}(e)^T h + \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}$$

$$c. \quad \nabla f_{\alpha}(e + \lambda h) = \nabla f_{\alpha}(e) + \lambda h + o(\lambda h)$$

$$\text{où } o(\lambda h) \in \mathcal{R}^n \text{ et } \|o(\lambda h)\| \leq \frac{\lambda^2}{1 - |\lambda|}.$$

Démonstration

a. Nous partons de la définition de la fonction de pénalité intérieure évaluée au point $e + \lambda h$:

$$\begin{aligned} f_{\alpha}(e + \lambda h) &= \alpha c^T (e + \lambda h) + p(e + \lambda h) \\ &= \alpha c^T e + \alpha c^T \lambda h + p(e + \lambda h). \end{aligned}$$

En vertu de la proposition 2.2.2, point a, nous avons la minoration

$$f_{\alpha}(e + \lambda h) \geq f_{\alpha}(e) + \alpha c^T \lambda h - \lambda e^T h.$$

Par la mise en évidence du vecteur λh , nous obtenons:

$$f_{\alpha}(e + \lambda h) \geq f_{\alpha}(e) + \lambda (\alpha c^T - e^T) h$$

ou de manière équivalente

$$f_{\alpha}(e + \lambda h) \geq f_{\alpha}(e) + \lambda \nabla f_{\alpha}(e)^T h.$$

b. En appliquant le point b. de la proposition 2.2.2. , nous avons les relations suivantes :

$$\begin{aligned} f_{\alpha}(e + \lambda h) &= \alpha c^T (e + \lambda h) + p(e + \lambda h) \\ &= \alpha c^T e + \alpha c^T \lambda h + p(e + \lambda h) \\ &\leq f_{\alpha}(e) + \alpha c^T \lambda h - \lambda e^T h + \frac{\|\lambda h\|^2}{2} \frac{1}{1 - \|\lambda h\|_{\infty}} \end{aligned}$$

Nous pouvons transformer la dernière inégalité en considérant les trois expressions suivantes qui découlent de l'hypothèse $\|h\| = 1$:

$$\begin{aligned} \text{i.} \quad & \|\lambda h\|^2 = \lambda^2 \\ \text{ii.} \quad & \|\lambda h\|_{\infty} = |\lambda| \\ \text{iii.} \quad & \frac{1}{1 - \|\lambda h\|_{\infty}} = \frac{1}{1 - |\lambda|} \end{aligned} \quad (26)$$

Nous obtenons alors

$$f_{\alpha}(e + \lambda h) \leq f_{\alpha}(e) + \lambda \nabla f_{\alpha}(e)^T h + \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}.$$

c. En utilisant la formule de dérivation de la fonction f_{α} , il vient que

$$\nabla f_{\alpha}(e + \lambda h) = \alpha c + \nabla p(e + \lambda h)$$

et, par le point c. de la proposition 2.2.2. , nous obtenons

$$\begin{aligned} \nabla f_{\alpha}(e + \lambda h) &= \alpha c + \nabla p(e) + \lambda h + o(\lambda h) \\ &= \nabla f_{\alpha}(e) + \lambda h + o(\lambda h) \end{aligned}$$

$$\text{où } o(\lambda h) \in \mathbb{R}^n \text{ et } \|o(\lambda h)\| \leq \frac{\|\lambda h\|^2}{1 - \|\lambda h\|_{\infty}}.$$

$$\text{C'est-à-dire } \|o(\lambda h)\| \leq \frac{\lambda^2}{1 - |\lambda|} \quad \text{par (26)} \quad \blacksquare$$

Note : Nous pouvons remarquer que cette proposition est valable pour toute fonction du type $f(x) = \mu c^T x + p(x) + k$ où k est une constante.

3.2. La fonction potentielle

La fonction potentielle est définie pour chaque paramètre $v \leq \hat{v}$ par

$$x \in \overset{\circ}{S} \rightarrow f_q(x, v) = q \log(c^T x - v) + p(x)$$

où $q \geq n$ est un nombre fixé.

Ses dérivées sont :

$$\nabla f_q(x, v) = \frac{q}{c^T x - v} c - x^{-1} \quad \nabla f_q(e, v) = \frac{q}{c^T e - v} c - e \quad (27)$$

$$\nabla^2 f_q(x, v) = \frac{-q}{(c^T x - v)^2} c c^T + X^{-2} \quad \nabla^2 f_q(e, v) = \frac{-q}{(c^T e - v)^2} c c^T + I \quad (28)$$

Propriété 1

La fonction potentielle est non convexe.

Démonstration

Rappelons qu'une fonction f est non convexe signifie

$$\exists x, y \in \mathfrak{R}^n \text{ tels que } f(y) < f(x) + f'_x(x)(y - x).$$

Dès lors, pour montrer la propriété, il suffit de trouver x, y dans $\overset{\circ}{S}$, c, v et q tels que

$$q \log(c^T y - v) + p(y) < q \log(c^T x - v) + p(x) + (y - x) \left(\frac{q c}{c^T x - v} - x^{-1} \right)^T$$

Posons

$$\begin{cases} x = e \\ y^T = (\frac{1}{2} \dots \frac{1}{2}) \\ c^T = (\frac{1}{n} \dots \frac{1}{n}) \\ q > n \\ v = 0 \end{cases}$$

En injectant ces valeurs dans l'inégalité, nous nous demandons maintenant si

$$q \log\left(\frac{1}{2}\right) - \sum_{i=1}^n \log \frac{1}{2} < q \log(1) + p(e) + \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \dots \\ \frac{1}{2} \end{bmatrix} [(q-n)/n \dots (q-n)/n]$$

ou de manière équivalente si

$$(q - n) \log \frac{1}{2} < (q - n) \left(-\frac{1}{2}\right).$$

Or $q > n$ et $\log \frac{1}{2} < -\frac{1}{2}$, la propriété est donc vérifiée. ■

Propriété 2

La fonction potentielle $x \rightarrow f_q(x, v)$ a un minimum unique pour tout v et $q \geq n$.

Avant de démontrer cette propriété, nous devons considérer la fonction potentielle multiplicative définie par :

$$F_q(x, v) = \frac{(c^T x - v)^q}{\prod_{i=1}^n x_i}$$

La relation entre les deux fonctions potentielles est donnée par

$$f_q(x, v) = \log F_q(x, v).$$

Nous établissons quelques résultats auxiliaires relatifs à cette nouvelle fonction, [6].

Lemme 3.2.1

Si $\overset{\circ}{S}$ est un ensemble borné et ζ un vecteur de \mathbb{R}^n
est tel que $\zeta \neq 0$ et $A \zeta = 0$,
Alors, il existe au moins deux indices j et k , $j \neq k$,
tels que $\zeta_j > 0$ et $\zeta_k < 0$.

Démonstration

Faisons une démonstration par l'absurde.
Nous pouvons envisager deux cas :

1. $\zeta \geq 0$

L'ensemble $\{x + \alpha \zeta \mid \alpha \geq 0\}$ est alors admissible car
 $A(x + \alpha \zeta) = Ax + \alpha A \zeta = b + \alpha 0 = b$ et $x + \alpha \zeta > 0$.

Mais ceci est contraire au fait que $\overset{\circ}{S}$ soit borné.

$$2. \zeta \leq 0$$

Nous pouvons montrer de la même manière que l'ensemble $\{x + \alpha \zeta \mid \alpha \leq 0\}$ est admissible.

Les deux cas mènent à une contradiction. De plus, il existe au moins deux indices, i et j , tels que ζ_i et ζ_j soient non nuls :

En effet, si $\exists ! i$ tel que $\zeta_i \neq 0$, alors nous avons $\zeta_i > 0$ ou $\zeta_i < 0$.

Ceci nous conduit à la même contradiction que précédemment et donc la thèse est démontrée. ■

Lemme 3.2.2

Pour tout $\zeta \in \mathbb{R}^n$ tel que $A \zeta = 0$ et $\zeta \neq 0$, nous avons
 $\zeta^T \nabla^2 F_q(x) \zeta > 0$

Démonstration

La thèse est équivalente à

$\zeta^T H \zeta > 0$ où H est la matrice hessienne de $F_q(x, v)$ divisée par $F_q(x, v)$ ($F_q(x, v) > 0$).

Développons l'expression de H :

$$\begin{aligned} H(x) &= \frac{1}{F_q(x, v)} \nabla^2 F_q(x, v) \\ &= \frac{1}{F_q(x, v)} \nabla (\nabla F_q(x, v)) \end{aligned}$$

Or, nous disposons de la relation

$$\frac{1}{F_q(x, v)} \nabla F_q(x, v) = \nabla \log F_q(x, v)$$

qui nous permet d'écrire les égalités suivantes :

$$\begin{aligned}
H(x) &= \frac{1}{F_q(x, v)} \nabla (\nabla \log F_q(x, v) \cdot F_q(x, v)) \\
&= \frac{1}{F_q(x, v)} [\nabla^2 \log F_q(x, v) \cdot F_q(x, v) + \nabla \log F_q(x, v) \cdot \nabla F_q(x, v)] \\
&= \nabla^2 \log F_q(x, v) + \left(\frac{1}{F_q(x, v)} \nabla F_q(x, v) \right) \left(\frac{1}{F_q(x, v)} \nabla F_q(x, v) \right)^T \\
&= \frac{-q}{(c^T x - v)^2} c c^T + X^{-2} + \left(\frac{q}{c^T x - v} c - x^{-1} \right) \left(\frac{q}{c^T x - v} c - x^{-1} \right)^T.
\end{aligned}$$

Dès lors, la forme quadratique $\zeta^T H(x) \zeta$ pour $\zeta \in \mathbb{R}^n$, $\zeta \neq 0$, peut s'exprimer comme suit :

$$\zeta^T H \zeta = \frac{\zeta^T c c^T \zeta}{(c^T x - v)^2} (-q) + \zeta^T X^{-2} \zeta + \zeta^T \left(\frac{q}{c^T x - v} c - x^{-1} \right) \left(\frac{q}{c^T x - v} c - x^{-1} \right)^T \zeta$$

En posant $\gamma = \frac{c^T \zeta}{c^T x - v}$ et après quelques opérations élémentaires, nous obtenons

$$\zeta^T H \zeta = q(q-1)\gamma^2 - 2q x^{-T} \zeta \gamma + (x^{-T} \zeta)^2 + \zeta^T X^{-2} \zeta.$$

La forme quadratique est donc un polynôme de degré deux en γ .

Par conséquent, montrer que la forme quadratique est strictement positive revient à montrer que le discriminant du polynôme est strictement négatif.

Appelons D le discriminant divisé par quatre :

$$D = q^2 (x^{-T} \zeta)^2 - q(q-1) [(x^{-T} \zeta)^2 + \zeta^T X^{-2} \zeta]$$

Après simplification, nous obtenons :

$$D = q [(x^{-T} \zeta)^2 - (q-1) \zeta^T X^{-2} \zeta]$$

ou encore,

$$D = q \left[\left(\sum_{i=1}^n \frac{\zeta_i}{x_i} \right)^2 - (q-1) \sum_{i=1}^n \left(\frac{\zeta_i}{x_i} \right)^2 \right].$$

Essayons de décomposer l'expression de D pour examiner plus facilement son signe :
Par le lemme 3.2.1., nous pouvons affirmer qu'il existe j et k tels que

$$\frac{\zeta_j}{x_j} > 0 \text{ et } \frac{\zeta_k}{x_k} < 0 \quad (29)$$

En outre, nous pouvons définir l'indice l de la manière suivante :

$$l = k \text{ si } \sum_{i=1}^n \frac{\zeta_i}{x_i} > 0 \text{ et } l = j \text{ si } \sum_{i=1}^n \frac{\zeta_i}{x_i} \leq 0$$

Nous pouvons en déduire que

$$\frac{\zeta_l}{x_l} \sum_{i \neq l} \frac{\zeta_i}{x_i} < 0. \quad (30)$$

En effet, considérons les deux possibilités :

i. $l = k$

Nous avons alors $\frac{\zeta_l}{x_l} < 0$ par (29) et

$$\sum_{i \neq l} \frac{\zeta_i}{x_i} = \sum_{i=1}^n \frac{\zeta_i}{x_i} - \frac{\zeta_l}{x_l} > 0.$$

D'où le produit est strictement négatif.

ii. $l = j$

Dans ce cas, $\frac{\zeta_l}{x_l} > 0$ par (29) et $\sum_{i \neq l} \frac{\zeta_i}{x_i} < 0$.

Le produit est donc aussi strictement négatif.

Finalement, l'expression de D est donné par

$$D = q \left\{ \left(\sum_{i \neq l} \frac{\zeta_i}{x_i} \right)^2 - (q-1) \sum_{i \neq l} \left(\frac{\zeta_i}{x_i} \right)^2 \right\} + 2 \frac{\zeta_l}{x_l} \sum_{i \neq l} \frac{\zeta_i}{x_i} - (q-2) \left(\frac{\zeta_l}{x_l} \right)^2 \}.$$

Examinons le signe de chaque terme :

$$1. \quad 2 \frac{\zeta_l}{x_l} \sum_{i \neq l} \frac{\zeta_i}{x_i} < 0 \quad \text{par (30)}$$

$$2. \quad - (q-2) \left(\frac{\zeta_l}{x_l} \right)^2 < 0$$

$$3. \left(\sum_{i \neq 1} \frac{\xi_i}{x_i} \right)^2 - (q - 1) \sum_{i \neq 1} \left(\frac{\xi_i}{x_i} \right)^2 < 0$$

En effet, nous avons l'inégalité

$$\left| \sum_{i \neq 1} \frac{\xi_i}{x_i} \right| \leq \sum_{i \neq 1} \left| \frac{\xi_i}{x_i} \right|$$

" la valeur absolue d'une somme est inférieure ou égale à la somme des valeurs absolues ."

En élevant cette inégalité au carré, nous obtenons les relations suivantes :

$$\begin{aligned} \left(\sum_{i \neq 1} \frac{\xi_i}{x_i} \right)^2 &\leq \left(\sum_{i \neq 1} \left| \frac{\xi_i}{x_i} \right| \right)^2 \\ &\leq (n - 1) \sum_{i \neq 1} \frac{\xi_i^2}{x_i^2} \quad (\text{propriété des normes 1 et 2}) \\ &\leq (q - 1) \sum_{i \neq 1} \left(\frac{\xi_i}{x_i} \right)^2 \quad (q \geq n) \end{aligned}$$

Nous pouvons en conclure que D est strictement négatif, d'où la thèse. ■

Lemme 3.2.3

Si $\overset{\circ}{S} = \{ x \mid Ax = b, x > 0 \}$ est borné,

Alors, étant donné $\bar{x} \in \overset{\circ}{S}$, nous pouvons montrer

$$T = \left\{ x \in \overset{\circ}{S} \mid \frac{(c^T x - v)^q}{\prod_{i=1}^n x_i} \leq \frac{(c^T \bar{x} - v)^q}{\prod_{i=1}^n \bar{x}_i} \right\}$$

est borné et fermé.

Démonstration

- T est borné

Par hypothèse, nous savons que $\overset{\circ}{S}$ est borné.

Or $T \subseteq \overset{\circ}{S}$, donc nous pouvons affirmer que T est borné.

- T est fermé

Prenons une suite de points dans T, (x^k) , convergente vers un élément x.
Montrons que ce dernier appartient à T.

Autrement dit, vérifions que

$$\frac{(c^T x - v)^q}{\prod_{i=1}^n x_i} \leq \frac{(c^T \bar{x} - v)^q}{\prod_{i=1}^n \bar{x}_i}.$$

La suite (x^k) appartient à T, donc la suite $\left(\frac{(c^T x^k - v)^q}{\prod_{i=1}^n x_i^k}\right)$ est une suite bornée.

Par passage à la limite, lorsque k tend vers l'infini, nous obtenons :

$$(c^T x^k - v)^q \rightarrow (c^T x - v)^q \in \Re$$

$$\text{et } \prod_{i=1}^n x_i^k \rightarrow \prod_{i=1}^n x_i \neq 0$$

Comme $\prod_{i=1}^n x_i \neq 0$, nous avons $x > 0$ (31)

$$\text{et } \frac{(c^T x - v)^q}{\prod_{i=1}^n x_i} \leq \frac{(c^T \bar{x} - v)^q}{\prod_{i=1}^n \bar{x}_i}$$

De plus, chaque élément de la suite (x^k) appartient à T, d'où $A x^k = b \quad \forall k$.
Par passage à la limite, quand k tend vers l'infini, nous avons

$$A x = b \quad (32)$$

Par (31) et (32), nous obtenons $x \in \overset{\circ}{S}$.

Nous avons ainsi vérifié les conditions d'appartenance à T pour x. ■

Proposition 3.2.4

La fonction potentielle multiplicative est strictement convexe sur $\overset{\circ}{S}$.

Démonstration

Par définition d'une fonction strictement convexe, nous devons montrer que

$$Fq(y, v) > Fq(x, v) + \nabla Fq(x, v)^T (y - x) \quad \forall x, y \in \overset{\circ}{S} \text{ tels que } y \neq x.$$

Sans perte de généralité, supposons $y > x$.

Comme la fonction potentielle multiplicative est différentiable, nous pouvons prendre son développement de Taylor d'ordre deux :

$$\begin{aligned} Fq(y, v) &= Fq(x, v) + \nabla Fq(x, v)^T (y - x) \\ &\quad + 1/2 (y - x)^T \nabla^2 Fq(\bar{x}, v) (y - x) \end{aligned}$$

$$\text{où } \bar{x} = x + \Theta (y - x) \text{ avec } 0 < \Theta < 1.$$

Nous avons $\bar{x} \in \overset{\circ}{S}$ car

$$\begin{aligned} A \bar{x} &= A x + \Theta A (y - x) \\ &= b + \Theta (A y - A x) \\ &= b \quad (y \text{ et } x \in \overset{\circ}{S}) \end{aligned}$$

et de plus,

$$\begin{aligned} \bar{x} &= x + \Theta (y - x) > 0 \\ (x, y > 0; 0 < \Theta < 1; y > x) \end{aligned}$$

La thèse s'écrit maintenant

$$\frac{1}{2} (y - x)^T \nabla^2 Fq(\bar{x}) (y - x) > 0.$$

Elle est vérifiée en appliquant le lemme 3.2.2. avec $\zeta = y - x$

où $A \zeta = A (y - x) = 0$ et $\zeta = y - x \neq 0$. ■

Proposition 3.2.5

La fonction potentielle multiplicative possède un minimum dans $\overset{\circ}{S}$.

Démonstration

La thèse revient à prouver l'existence d'une solution au problème

$$(P) \equiv \left[\begin{array}{l} \text{Minimiser} \quad \frac{(c^T x - v)^q}{\prod_{i=1}^n x_i} \\ \text{sous} \quad \begin{array}{l} Ax = b \\ x > 0 \end{array} \end{array} \right.$$

$$\text{Soit } T = \left\{ x \in \overset{\circ}{S} \mid \frac{(c^T x - v)^q}{\prod_{i=1}^n x_i} \leq \frac{(c^T \bar{x} - v)^q}{\prod_{i=1}^n \bar{x}_i} \right\}$$

où $\bar{x} \in \overset{\circ}{S}$.

Par le lemme 3.2.3, T est borné et fermé.

Or une fonction continue, $x \rightarrow \frac{(c^T x - v)^q}{\prod_{i=1}^n x_i}$, sur un borné fermé, T , atteint sa borne inférieure; soit x^* cette borne.

Nous pouvons montrer que x^* est solution du problème (P) car

si $x \notin T$ et $x \in \overset{\circ}{S}$, nous avons la relation suivante :

$$\frac{(c^T x - v)^q}{\prod_{i=1}^n x_i} > \frac{(c^T \bar{x} - v)^q}{\prod_{i=1}^n \bar{x}_i} \geq \frac{(c^T x^* - v)^q}{\prod_{i=1}^n x_i^*}.$$

La thèse est ainsi montrée. ■

Démonstration de la propriété 2

Nous savons par la proposition 3.2.5 que $F_q(x, v)$ possède un minimum dans $\overset{\circ}{S}$.
En vertu de la proposition 3.2.4., nous pouvons assurer l'unicité du minimum pour la fonction potentielle multiplicative.

Or nous disposons de la relation entre les deux fonctions potentielles :

$$\nabla f_q(x, v) = \frac{\nabla F_q(x, v)}{F_q(x, v)}$$

Il suit que si x^* est l'unique minimum de $F_q(x, v)$, x^* est l'unique point stationnaire de $f_q(x, v)$.

En outre,

$$\frac{1}{F_q(x, v)} \nabla^2 F_q(x, v) = \nabla^2 f_q(x, v) + \nabla f_q(x, v) \nabla f_q(x, v)^T.$$

Évaluons cette expression en x^* . Nous obtenons :

$$\frac{1}{F_q(x^*, v)} \nabla^2 F_q(x^*, v) = \nabla^2 f_q(x^*, v).$$

Comme $\nabla^2 F_q(x^*, v)$ est définie positive et $F_q(x^*, v)$ est strictement positif, nous obtenons finalement que $\nabla^2 f_q(x^*, v)$ est définie positive au point stationnaire x^* .
Il est donc le minimum unique de $f_q(x, v)$. ■

Propriété 3

La fonction potentielle $f_q(\bullet, v)$ est unimodale le long de toute direction admissible.

Démonstration

Soient $x \in \overset{\circ}{S}$ et d une direction admissible.

Nous savons que F_q est strictement convexe et donc strictement quasi convexe sur l'intervalle $[x, x + \lambda d] \subseteq \overset{\circ}{S}$.

Nous nous référons alors au théorème 8.7.2 (voir [9], p 275) nous assurant que dans ce cas, F_q est strictement unimodale sur $[x, x + \lambda d]$.

En outre, nous avons la relation $f_q(x, v) = \log F_q(x, v)$ et la fonction logarithmique est strictement croissante.

Par conséquent,

$f_q(\cdot, v)$ est strictement unimodale le long de toute direction admissible. ■

Remarque

- Une fonction $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ est strictement unimodale sur $[a, b]$ s'il existe $t^* \in [a, b]$ tel que

$$f(t^*) = \min \{ f(t) \mid t \in [a, b] \}$$
 et si pour tout t_1, t_2 vérifiant la relation $a \leq t_1 < t_2 \leq b$, nous avons

$$t_2 \leq t^* \text{ implique que } f(t_1) > f(t_2)$$
 et $t^* \leq t_1 \text{ implique que } f(t_2) > f(t_1)$.
- Une fonction strictement unimodale sur $[a, b]$ a un minimum unique global sur $[a, b]$.

4. La trajectoire centrale

4.1. Définition de la trajectoire centrale

Définition 4.1.1

Le point central $x(\rho)$ associé au paramètre $\rho \in \mathbb{R}^+$ est défini comme :

$$x(\rho) = \operatorname{argmin} \{ f_\rho(x) \mid Ax = b, x > 0 \}$$

où f_ρ est une fonction auxiliaire.

Définition 4.1.2

La trajectoire centrale est la courbe

$$\rho \in \mathbb{R}^+ \rightarrow x(\rho) \in \overset{\circ}{S}$$

La trajectoire centrale comprend l'ensemble des optimums pour différentes fonctions auxiliaires basées sur la fonction barrière logarithmique.

Chaque fonction auxiliaire donne lieu à un paramétrage différent $\rho \rightarrow x(\rho)$ de la trajectoire.

Nous pouvons donner une interprétation géométrique de la trajectoire centrale paramétrée par α , le paramètre de la fonction de pénalité intérieure.

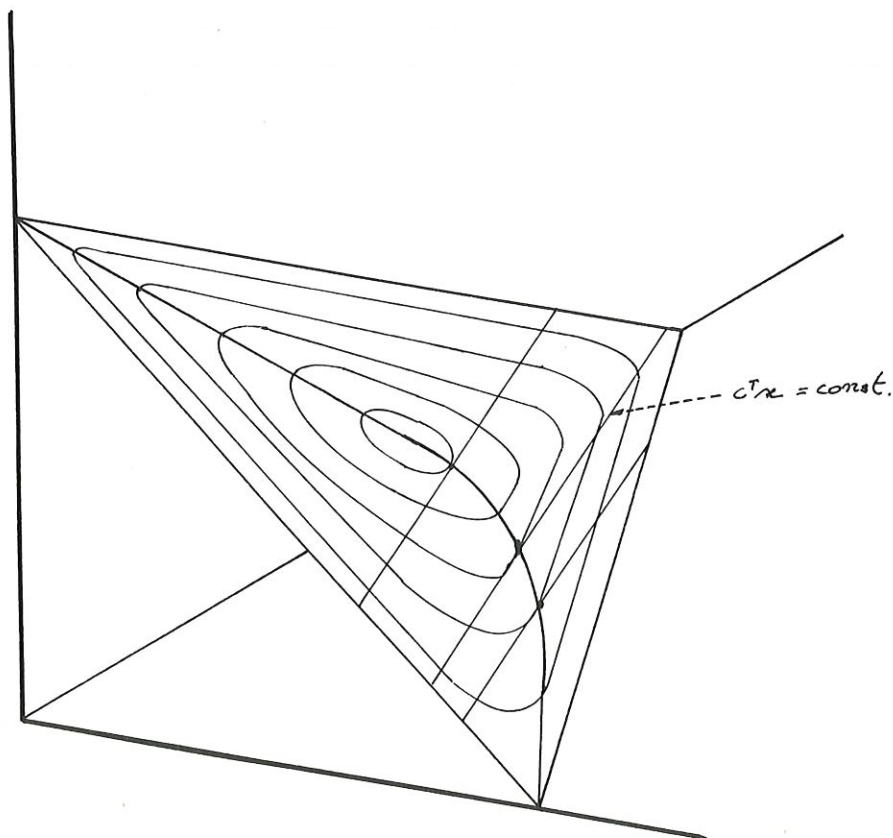
Considérons un point central $x(\alpha) = \operatorname{argmin}_{x \in S} \{\alpha c^T x + p(x)\}$

Ce point résout aussi le problème obtenu en contraignant le coût à sa valeur en $x(\alpha)$, c'est-à-dire,

$$x(\alpha) = \operatorname{argmin}_{x \in S} \{p(x) \mid c^T x = c^T x(\alpha)\}$$

Ce problème décrit le centre analytique de la fonction barrière p , sur une coupe de "coût constant" de l'ensemble admissible S .

Ceci est illustré par la figure ci-dessous.



4.2. Les paramétrages de la trajectoire centrale

4.2.1. Paramétrage primal-dual

Le paramétrage primal-dual est celui utilisé dans la définition des points centraux pour la fonction de pénalité intérieure.

$$x \rightarrow f_{\alpha}(x) = \alpha c^T x + p(x)$$

La dénomination "primal-dual" est due à la relation étroite entre le paramètre α et le saut de dualité associé à un point central $\Delta(\alpha)$.

Ceci est montré dans la proposition 4.2.1.2.

Cette proposition associe à chaque point central $x(\alpha)$, un point dual admissible $z(\alpha)$ et en conséquence un saut de dualité $\Delta(\alpha)$.

Lemme 4.2.1.1

Soit le problème

$$(P) \equiv \begin{cases} \text{Min } f(x) \\ \text{sous } Ax = b \\ x > 0 \end{cases}$$

Si x^ est solution de (P), alors*

$$Ax^* = b \text{ et } P_A \nabla f(x^*) = 0$$

Démonstration

Les conditions nécessaires d'optimalité, nous permettent d'affirmer :

$$x^* \text{ est solution de (P) implique qu'il existe } \lambda \in \mathbb{R}^m \text{ tel que} \\ \nabla f(x^*) + A^T \lambda = 0 \text{ et } Ax^* = b.$$

Recherchons l'expression de λ :

Nous partons de l'équation

$$A^T \lambda = -\nabla f(x^*).$$

En multipliant les deux membres par la matrice A , il vient que

$$AA^T \lambda = -A \nabla f(x^*).$$

Comme la matrice A est de rang plein, AA^T est inversible.

Nous obtenons ainsi :

$$\lambda = - (A A^T)^{-1} A \nabla f(x^*).$$

La condition d'optimalité devient alors :

$$\nabla f(x^*) - A^T (A A^T)^{-1} A \nabla f(x^*) = 0 \quad \text{et} \quad A x^* = b$$

ou encore

$$(I - A^T (A A^T)^{-1} A) \nabla f(x^*) = 0 \quad \text{et} \quad A x^* = b$$

c'est-à-dire

$$P_A \nabla f(x^*) = 0 \quad \text{et} \quad A x^* = b \quad \blacksquare$$

Proposition 4.2.1.2

Soit $x(\alpha)$ le point central associé à $\alpha > 0$.

Alors, $z(\alpha) = \frac{x(\alpha)^{-1}}{\alpha}$ est un point dual admissible et le saut de dualité

associé à $x(\alpha)$ et $z(\alpha)$ est $x(\alpha)^T z(\alpha) = \frac{n}{\alpha}$

Démonstration

Nous avons, par définition de la fonction de pénalité intérieure

$$f_\alpha(x) = \alpha c^T x + p(x)$$

et par dérivation,

$$\nabla f_\alpha(x) = \alpha c - x^{-1}.$$

De plus, par hypothèse, $x(\alpha)$ est le point central associé à $\alpha > 0$.

En vertu du lemme 4.2.1.1, nous pouvons affirmer :

$$P_A(\alpha c - x(\alpha)^{-1}) = 0$$

Comme l'opérateur projection est linéaire, nous obtenons

$$P_A \alpha c = P_A x(\alpha)^{-1} \quad \text{ou} \quad P_A (x(\alpha))^{-1} = \alpha c_p.$$

Puisque $x(\alpha) > 0$, $\frac{x(\alpha)^{-1}}{\alpha}$ est un point dual admissible en vertu de la proposition 1.3.2.

De plus, le saut de dualité est donné par

$$\frac{1}{\alpha} x(\alpha) x(\alpha)^{-T} = \frac{n}{\alpha} \quad \blacksquare$$

4.2.2. Paramétrage dual

Le paramétrage dual est celui associé à la fonction potentielle.

Il est appelé "dual" car le paramètre v est une borne inférieure à la valeur optimale \hat{v} et est toujours associé au coût dual $b^T w$ d'une solution duale admissible.

Le point central $x(v)$ est caractérisé uniquement par

$$\frac{q}{c^T x - v} c_p - P x^{-1} = 0 \quad \text{ou par} \quad \frac{P x^{-1}}{q / c^T x - v} = P c.$$

Le saut de dualité associé au point central $x(v)$ est donné par

$$\Delta(v) = \frac{n}{q} (c^T x - v).$$

4.3. Identification des trajectoires caractérisées par les deux paramétrages

Lemme 4.3.1

Les deux assertions suivantes sont équivalentes:

i) $x(\rho)$ est la solution du problème

$$(P) \equiv \begin{cases} \text{Minimiser} & f_\rho(x) \\ \text{sous} & A x = b \\ & x > 0 \end{cases}$$

ii) $\bar{x}(\rho)$ est la solution du problème (\bar{P}) , dans l'espace transformé par un "scaling".
où

$$(\bar{P}) \equiv \begin{cases} \text{Minimiser} & \bar{f}_\rho(y) \\ \text{sous} & \bar{A} y = b \\ & y > 0 \end{cases}$$

Démonstration :

Considérons la transformation "scaling" suivante :

$$D^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$x \rightarrow D^{-1} x = y$$

où D est une matrice diagonale positive.

Soit $x(\rho)$ la solution du problème (P) .

Notons $\bar{x}(\rho) = D^{-1} x(\rho)$ et montrons que $\bar{x}(\rho)$ est la solution du problème (\bar{P}) .

Pour cela, nous devons vérifier que $\bar{x}(\rho)$ satisfait les deux conditions suivantes :

1) $\bar{x}(\rho)$ est admissible pour le problème (\bar{P}) :

- $\bar{A}(\bar{x}(\rho)) = A D D^{-1} x(\rho) = b$
- $\bar{x}(\rho) = D^{-1} x(\rho) > 0$ car $x(\rho)$ est admissible pour (P) et D est une matrice diagonale positive.

2) $\bar{f}_\rho(\bar{x}(\rho)) \leq \bar{f}_\rho(y) \quad \forall y > 0$ vérifiant $\bar{A} y = b$.

Il est équivalent de montrer que

$$f_\rho(D \bar{x}(\rho)) \leq f_\rho(D y) \quad \forall y > 0 \text{ tel que } \bar{A} y = b,$$

ou encore

$$f_\rho(x(\rho)) \leq f_\rho(D y) \quad \forall y > 0 \text{ tel que } \bar{A} y = b.$$

Cette dernière assertion est vérifiée puisque $x(\rho)$ est solution du problème (P) et Dy est admissible ($A D y = \bar{A} y = b$ et $D y > 0$).

L'implication dans l'autre sens se démontre de manière tout à fait analogue. ■

La proposition suivante nous montre l'équivalence des deux trajectoires centrales . L'une paramétrée par α , le paramètre de la fonction de pénalité intérieure; l'autre par v , le paramètre de la fonction potentielle.

Proposition 4.3.2

Nous pouvons établir les deux assertions suivantes :

$$i. \quad \forall \alpha > 0 \quad \exists v \leq \hat{v} \text{ tel que } x(\alpha) = x(v)$$

$$ii. \quad \forall v \geq \hat{v} \quad \exists \alpha > 0 \text{ tel que } x(v) = x(\alpha)$$

Démonstration

i. Soit $\alpha > 0$.

$x(\alpha)$ est le point central associé au paramètre α . Il est, par définition, la solution du problème

$$\left[\begin{array}{ll} \text{Minimiser} & f_{\alpha}(x) \\ \text{sous} & A x = b \\ & x > 0 \end{array} \right.$$

Considérons la transformation "scaling" :

$$\begin{aligned} X^{-1} : \mathcal{R}^n &\rightarrow \mathcal{R}^n \\ x &\rightarrow X^{-1} x \end{aligned}$$

où $X = \text{diag}(x(\alpha))$

En vertu du lemme 4.3.1, dans l'espace transformé, $e = X^{-1} x(\alpha)$ est la solution du problème

$$\left[\begin{array}{ll} \text{Minimiser} & \bar{f}_{\alpha}(y) \\ \text{sous} & A X y = b \\ & y > 0 \end{array} \right.$$

Par le lemme 4.2.1.1, nous pouvons affirmer que

$$P_{AX} \nabla \bar{f}_{\alpha}(e) = 0.$$

Développons cette équation en considérant

$$\bar{f}_{\alpha}(y) = f_{\alpha}(X y) = \alpha (X c)^T y - \sum_{i=1}^n \ln y_i - \sum_{i=1}^n \ln x_i(\alpha).$$

Dès lors, nous avons

$$P_{AX} [\alpha X c - e] = 0$$

ou encore

$$\boxed{\alpha P_{AX}(X c) - P_{AX} e = 0} \quad (33)$$

par linéarité de l'opérateur projection.

Notre problème consiste en fait à trouver v de sorte que e soit aussi solution du problème:

$$\left[\begin{array}{ll} \text{Minimiser} & \bar{f}_q(y, v) \\ \text{sous} & A X y = b \end{array} \right.$$

Supposons que le point e soit solution de ce dernier problème.

Le lemme 4.2.1.1 nous assure que

$$P_{AX} \nabla \bar{f}_q(e, v) = 0.$$

En considérant les égalités suivantes,

$$\begin{aligned} \bar{f}_q(y, v) &= f_q(X y, v) \\ &= q \ln((X c)^T y - v) + p(y) + p(x(\alpha)) \end{aligned}$$

le paramètre v doit vérifier l'équation

$$P_{AX} [(q / (X c)^T e - v) X c - e] = 0.$$

c'est-à-dire

$$\frac{q}{(X c)^T e - v} P_{AX} (X c) - P_{AX} e = 0 \quad (34)$$

L'identification des expressions (33) et (34), nous donne

$$\alpha = \frac{q}{(X c)^T e - v} = \frac{q}{c^T x(\alpha) - v}$$

Et par conséquent, nous obtenons la valeur du paramètre v

$$v = c^T x(\alpha) - \frac{q}{\alpha}.$$

La première assertion est ainsi vérifiée.

ii. Soit $v \leq \hat{v}$.

$x(v)$ est la solution du problème

$$\left[\begin{array}{ll} \text{Minimiser} & f_q(x, v) \\ \text{sous} & A x = b \\ & x > 0 \end{array} \right.$$

Considérons la transformation "scaling" :

$$X'^{-1} : \mathcal{R}^n \rightarrow \mathcal{R}^n$$

$$x \rightarrow X'^{-1} x$$

$$\text{où } X' = \text{diag} (x(v))$$

Dans l'espace transformé, e est la solution du problème

$$\left[\begin{array}{ll} \text{Minimiser} & \bar{f}_q(y, v) \\ \text{sous} & A X' y = b \\ & y > 0 \end{array} \right.$$

Par conséquent, nous avons :

$$P_{AX'} \cdot \nabla \bar{f}_q(e, v) = 0$$

c'est-à-dire

$$\boxed{\frac{q}{(X' c)^T e - v} P_{AX'} (X' c) - P_{AX'} e = 0} \quad (35)$$

Nous devons trouver le paramètre α tel que e soit aussi solution du problème

$$\left[\begin{array}{ll} \text{Minimiser} & \bar{f}_\alpha(y) \\ \text{sous} & A X' y = b \\ & y > 0 \end{array} \right.$$

Le paramètre α doit donc satisfaire l'équation suivante

$$P_{AX'} \cdot \nabla \bar{f}_\alpha(e) = 0$$

c'est-à-dire

$$\boxed{\alpha P_{AX'} (X' c) - P_{AX'} e = 0} \quad (36)$$

En égalant les expressions (35) et (36), nous obtenons :

$$\alpha = \frac{q}{c^T x(v) - v}$$

Ainsi se termine la preuve. ■

Chapitre IV

De l'algorithme conceptuel à l'algorithme pouvant être mis en œuvre

Dans ce chapitre, nous décrivons le cheminement général depuis l'algorithme conceptuel jusqu'à l'algorithme que l'on peut mettre en œuvre c'est-à-dire qui comprend les éléments nécessaires à la convergence vers une solution optimale du problème (PL).

1. L'algorithme conceptuel

Soit $\rho \in (\omega^-, \omega^+) \rightarrow x(\rho)$ un paramétrage de la trajectoire centrale, avec $\omega^+ > \omega^-$ où ω^+ peut être l'infini.

La trajectoire centrale convergera vers une solution optimale du problème (PL), par une variation du paramètre ρ .

L'algorithme conceptuel suit simplement la trajectoire centrale paramétrée par ρ selon le modèle suivant :

Algorithme 1.1 " L'algorithme conceptuel "

Soit $x^\circ \in \mathring{S}$,

$\rho_0 \in (\omega^-, \omega^+)$ tel que $x^\circ = x(\rho_0)$

$k = 0$

Répéter jusqu'à convergence

- Choisir $\rho_{k+1} > \rho_k$
- Appeler un algorithme de minimisation interne pour trouver $x^{k+1} = x(\rho_{k+1})$.
- $k = k + 1$

Ce modèle génère simplement une suite de points centraux. En réalité, les algorithmes dépendront du paramétrage, de l'initialisation, (choix de ρ_0 et de x°), et surtout du critère de mise à jour du paramètre.

Le paramétrage et le critère de mise à jour diffèrent suivant la fonction auxiliaire et donc donnent lieu à des algorithmes différents.

2. L'algorithme de minimisation interne

L'algorithme interne est essentiellement le même pour toutes les méthodes. Nous utilisons l'algorithme S-S-D [cfr chap II]. L'algorithme conceptuel peut donc être complété comme suit:

Algorithme 2.1

Soit $x^0 \in \overset{\circ}{S}$

$f : \overset{\circ}{S} \rightarrow \Re$ continûment différentiable

$k = 0$

Répéter jusqu'à convergence

- Choisir $\rho_{k+1} > \rho_k$
- $j = 0$
- $y^0 = x^k$

Répéter jusqu'à convergence

- "Scaling ": $\bar{A} = A Y_j$
 $\bar{g} = Y_j \nabla f(y^j)$
- Direction : $\bar{h} = -P_{\bar{A}} \bar{g}$
- Recherche linéaire :

$$\begin{aligned}\bar{z} &= e + \lambda \bar{h} \\ \bar{z} &> 0\end{aligned}$$

- "Scaling" inverse : $y^{j+1} = Y_j \bar{z}$
- $j = j + 1$
- $k = k + 1$

Le point important que nous devons discuter est le critère d'arrêt de l'algorithme de minimisation interne.

Il est impossible de trouver exactement un point central en un temps fini et nous voulons construire des algorithmes polynomiaux. Aussi, d'un point de vue pratique, l'algorithme interne devrait se terminer aussi vite que possible.

Nous devons alors renoncer à la détermination de points centraux et travailler près de la trajectoire centrale.

Des critères précis doivent être définis pour considérer un point "proche" de la trajectoire centrale.

2.1. Le critère de proximité

Comme il est impossible de calculer un point central $x(\rho)$ exactement, nous avons besoin d'un critère pour décider quand un point x est "proche" du point central $x(\rho)$. Le critère d'arrêt de l'algorithme interne sera basé sur le critère de proximité.

Notons $h(x, \rho)$ la direction S-S-D dans l'espace d'origine et

$\bar{h}(x, \rho)$ la direction S-S-D dans l'espace transformé par un "scaling".

Définition 2.1.1

La longueur du pas S-S-D dans l'espace transformé,

$$\delta(x, \rho) = \|\bar{h}(x, \rho)\|$$

est la mesure de proximité d'un point x par rapport à $x(\rho)$.

Nous pouvons donner une interprétation géométrique de la mesure de proximité dans le cas particulier de la fonction barrière logarithmique :

Soit χ le centre analytique de S , c'est-à-dire l'unique point qui minimise la fonction barrière.

Supposons $\chi = e$.

Dans ce cas simple, la fonction barrière autour du point e peut s'écrire

$$p(e + h) = p(e) + \nabla p(e)^T h + \frac{1}{2} h^T \nabla^2 p(e) h + o(h)$$

ou encore puisque $p(e) = 0$ et $\nabla^2 p(e) = I$,

$$p(e + h) = \frac{1}{2} \|h\|^2 + o(h).$$

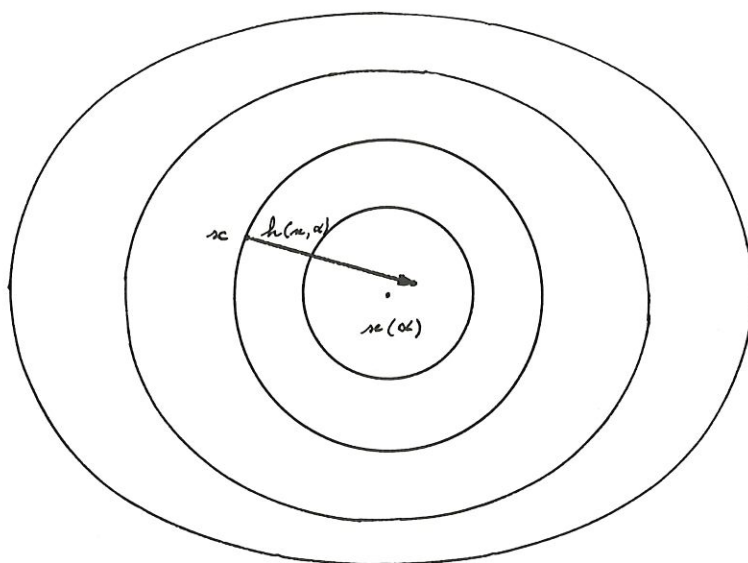


figure 2.1.1

La figure montre des courbes de niveau pour la fonction

$$p(e + h) = \frac{1}{2} \|h\|^2 + o(h).$$

Pour des petites valeurs de $\|h\|$ (petites par rapport à 1), les courbes de niveau sont presque sphériques.

(conséquence du comportement de la fonction logarithmique autour de 1). En d'autres termes, l'approximation quadratique est très bonne près du centre.

Deux grandeurs semblent être de bonnes mesures de proximité par rapport au centre e :

$$\|x - e\| \text{ et } p(x) - p(e).$$

Si ces mesures sont inférieures à ε où $0 < \varepsilon \leq 1$, x est dans une région où l'approximation quadratique est bonne.

Malheureusement, ces deux critères de proximité ne sont pas utiles pour des méthodes pratiques car leur calcul dépend de la connaissance du centre.

Décrivons le troisième et meilleur critère : le pas S-S-D, $h(x)$.
Comme l'illustre la figure, aux points où l'approximation quadratique est bonne, le pas S-S-D est une bonne approximation du pas vers le minimum de la fonction;
c'est-à-dire

$$h(x) \cong e - x \text{ ou } \|h(x)\| \cong \|x - x(\alpha)\|.$$

Comme tous ces critères sont similaires près de e , l'extension au cas général est évidente: précéder le calcul de proximité par un scaling.

Cette interprétation informelle de la mesure de proximité convient également pour la fonction de pénalité intérieure puisque celle-ci ne diffère de la fonction barrière logarithmique que par un terme linéaire. Rien ne change dans l'analyse de l'approximation du second ordre.

Une analyse formelle du critère de proximité est faite pour chacune des fonctions auxiliaires dans les chapitres V et VI.

Nous énonçons à présent le critère de proximité :

Si $\|\bar{h}(x)\| \ll 1$, nous considérons que x est proche du point central. La valeur de

la borne de $\|\bar{h}(x)\|$ sera précisée pour chaque fonction auxiliaire. Pour l'instant, considérons la borne égale à ε ($0 < \varepsilon \leq 1$).

Algorithme 2.1.2

Soit $x^\circ \in \overset{\circ}{S}$

$f: \overset{\circ}{S} \rightarrow \Re$

$k = 0$

Répéter jusqu'à convergence

- Choisir $\rho_{k+1} > \rho_k$
- Répéter jusqu'à ce que $\|\bar{h}(x, \rho_{k+1})\| < \varepsilon$
" Algorithme S-S-D "
- $k = k + 1$

3. Mise à jour du paramètre

Il existe deux types de mise à jour :

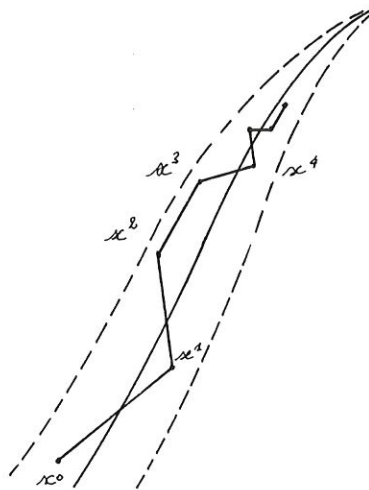
- Une mise à jour "à petits pas" :

Celle-ci force l'algorithme à suivre la trajectoire de telle façon que tous les points générés soient près de la trajectoire.

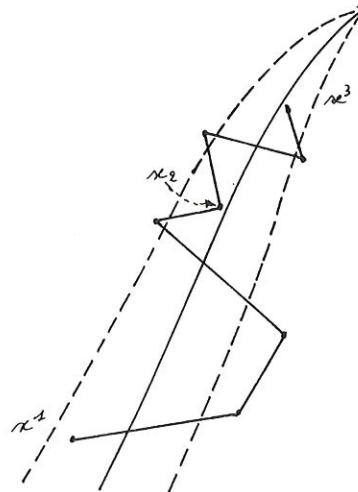
L'algorithme interne s'exécute habituellement en une itération par itération de l'algorithme principal.

- Une mise à jour "à grands pas" :

Elle utilise une grande variation du paramètre. Plusieurs itérations de l'algorithme interne sont donc nécessaires pour approcher le point central correspondant à ρ_{k+1} .



mise à jour
"à petits pas"



mise à jour
"à grands pas"

Nous considérons, dans ce texte, uniquement le deuxième type de mise à jour. La mise à jour a pour effet de réduire le saut de dualité d'un rapport $\beta \in (0, 1)$. Pour chaque fonction auxiliaire, nous aurons une procédure de mise à jour particulière du paramètre.

fonction pénalisée $f_{\alpha}(x)$:

$$\alpha_{k+1} := \frac{1}{\beta} \alpha_k$$

Ainsi le saut de dualité devient

$$\Delta(\alpha_{k+1}) = \beta \Delta(\alpha_k).$$

Nous avons bien $\Delta(\alpha_{k+1}) < \Delta(\alpha_k)$.

La réduction du saut de dualité obtenue en une itération de l'algorithme est égale à la réduction β .

fonction potentielle $f_q(x, v)$:

$$\text{Comme } \Delta(v) = \frac{n}{q} (c^T x - v),$$

cela signifie qu'étant donné une borne inférieure v , le point central $x(v)$ définit une nouvelle borne inférieure v' par :

$$c^T x - v' = \frac{n}{q} (c^T x - v).$$

Si $n = q$, il n'y a pas de réduction du saut.

Si $q > n$, alors toute valeur de v' telle que

$$c^T x - v' = \beta (c^T x - v) \text{ avec } \beta \in \left(\frac{n}{q}, 1\right)$$

fournit une nouvelle borne inférieure.

La valeur de q détermine la réduction maximale permise du saut par itération et en conséquence la longueur maximale du pas.

Lorsque la borne inférieure est mise à jour, il y a une réduction immédiate du saut.

En x^{k+1} , le saut sera encore inférieur car $c^T x^{k+1} < c^T x^k$.

La mise à jour est donc pessimiste puisque la réduction réelle du saut est meilleure que celle que nous visions.

Les procédures de mise à jour du paramètre ρ seront précisées dans les chapitres V et VI pour chacune des fonctions auxiliaires.

Nous pouvons à présent établir le tableau récapitulatif donnant les caractéristiques principales des deux paramétrages.

PARAMETRAGE		
	Primal-dual	dual
$f_p(x)$	$\alpha c^T x + p(x)$	$q \log(c^T x - v) + p(x)$
$\nabla f_p(x)$	$\alpha c - x^{-1}$	$\frac{q}{c^T x - v} c - x^{-1}$
$\Delta_p(x)$	$\frac{n}{\alpha}$	$\frac{n}{q} (c^T x - v)$
mise à jour du paramètre ρ	$\alpha' := \frac{\alpha}{\beta}$	$(c^T x - v') = \beta(c^T x - v)$
		$\beta \in (\frac{n}{q}, 1)$
		$v' = \beta v + (1 - \beta) c^T x$

tableau 3.1

4. Critère d'arrêt de l'algorithme principal

Comme nous pouvions nous y attendre, l'algorithme s'arrête quand le saut de dualité est petit.

Nous supposons qu'un point central initial est donné, $x^0 = x(\rho_0)$ et que les procédures de mise à jour se font selon le tableau 3.1.

Nous supposons aussi que $q = O(n)$, $q \geq n$.

Le critère d'arrêt est $\Delta(x^k, \rho_k) < 2^{-L}$

où L est la longueur des données du problème (PL) [cfr chapitre I].

Justifions ce critère d'arrêt :

Lemme 4.1

Si x est un sommet de l'ensemble admissible S , alors nous pouvons établir les deux assertions suivantes :

$$i) \text{ pour } i = 1, \dots, n \quad x_i = 0 \text{ ou } x_i > 2^{-L}$$

$$ii) c^T x - \hat{v} = 0 \text{ ou } c^T x - \hat{v} > 2^{-L}$$

Démonstration

Nous pouvons supposer s.p.d.g que $\hat{v} = 0$.

Le système $Ax = b$ peut être partitionné :

$$A_B x_B + A_N x_N = 0$$

où x_B est un vecteur fait des composantes positives de x ,

A_B est une matrice avec les colonnes correspondantes de A ,

$$x_N = 0$$

Nous obtenons ainsi $A_B x_B = b$.

Supposons que A_B soit une matrice carrée;

x est un sommet de l'ensemble admissible, donc x est une solution de base réalisable;

à x_B correspond une base A_B dont les colonnes sont linéairement indépendantes.

Ainsi, A_B est une matrice inversible.

Par conséquent,

$$x_B = A_B^{-1} b = \frac{(\text{cof } A_B)^T b}{\det A_B}$$

$$c^T x = \frac{c_B^T (\text{cof } A_B)^T b}{\det A_B}.$$

or $\det A_B < 2^L$ par définition de L et les numérateurs sont entiers, (les données du problème (PL) sont, par hypothèse, entières), d'où la thèse suit. ■

Le critère d'arrêt est précisément la deuxième assertion du lemme.

Si un algorithme obtient un point tel que $c^T \bar{x} - \hat{v} < 2^{-L}$, alors une solution optimale peut être obtenue par le lemme de purification (que nous ne démontrons pas), [4] et [8].

Lemme de purification

Etant donné $x \in S$, il existe une procédure qui calcule un sommet \bar{x} de S tel que $c^T \bar{x} \leq c^T x$ en au plus $O(n^3)$ opérations arithmétiques.

Si nous avons $c^T x - \hat{v} < 2^{-L}$ où $x \in S$, en vertu du lemme de purification, il existe un sommet \bar{x} tel que

$$c^T \bar{x} \leq c^T x.$$

Dès lors, nous avons $c^T \bar{x} - \hat{v} < 2^{-L}$.

Par le lemme 4.1., nous obtenons

$$c^T \bar{x} - \hat{v} = 0.$$

ou encore

$$c^T \bar{x} = \hat{v}.$$

Nous avons ainsi trouvé \bar{x} , une solution optimale.

5. L'algorithme pouvant être mis en oeuvre.

Nous disposons maintenant de tous les éléments nécessaires pour construire l'algorithme implémentable.

Algorithme 5.1

Soit $\varepsilon \in (0, 0.5]$,

$\beta \in (0, 1)$ (pour le paramétrage dual, $\beta \in (\frac{n}{q}, 1)$),

ρ_0 un paramètre initial admissible,

$x^\circ \in \overset{\circ}{S}$ tel que $\delta(x^\circ, \rho_0) < \varepsilon$.

$k = 0$

Répéter jusqu'à ce que $\Delta(x^k, \rho_k) < 2^{-L}$

- Calculer ρ_{k+1} par la procédure de mise à jour du tableau 3.1.
- Appeler l'algorithme S-S-D jusqu'à ce que
$$\delta(x^{k+1}, \rho_{k+1}) < \varepsilon.$$
- $k = k + 1.$

5.1 Complexité de l'algorithme

Lemme 5.1.1

Supposons que l'algorithme commence avec $\Delta(\rho_0) < 2^L$.

Le nombre d'itérations est alors borné par $O(L / \log(\frac{1}{\beta}))$.

Démonstration

Nous avons après k itérations et avant le critère d'arrêt, la relation suivante :

$$2^{-L} \leq \Delta(\rho_k) \leq \beta^k \Delta(\rho_0) < \beta^k 2^L.$$

Il suit que $-L \leq k \log_2 \beta + L$.

$$\text{Alors nous obtenons } k \leq \frac{2L}{\log_2(\frac{1}{\beta})} = O\left(\frac{L}{\log_2(\frac{1}{\beta})}\right). \quad \blacksquare$$

Le choix de β :

La complexité de l'algorithme dépend du nombre d'itérations de l'algorithme interne S-S-D nécessaires à chaque itération de l'algorithme principal. Ceci repose sur le critère de proximité par rapport aux points centraux.

Nous pouvons néanmoins donner, à ce stade, une évaluation générale de la borne de complexité, selon le choix de β :

- 1) β est indépendant de n :

La complexité de l'algorithme est alors $O(L)$ mais l'algorithme interne S-S-D peut demander un grand nombre d'itérations. Nous verrons que la meilleure complexité de l'algorithme est $O(nL)$ itérations.

- 2) β dépend de \sqrt{n} , tel que $\frac{1}{\beta} = 1 + \frac{\gamma}{\sqrt{n}}$.

Le nombre d'itérations de l'algorithme interne dépend de γ et le nombre global d'itérations S-S-D sera borné par $O(\gamma \sqrt{n} L)$.

Ces résultats seront obtenus et justifiés de manière plus précise dans les deux derniers chapitres.

Chapitre V

La méthode de la fonction de pénalité intérieure

1. Préliminaires

L'algorithme proposé dans ce chapitre est la méthode utilisant la fonction de pénalité intérieure comme fonction auxiliaire.

$$f_{\alpha}(x) = \alpha c^T x + p(x) \quad \text{où } x \in \overset{\circ}{S} \\ \alpha \in \Re$$

Cette méthode combine ainsi la fonction objectif du problème (PL) avec contraintes à la fonction barrière logarithmique et réalise une suite de minimisations sans contrainte de cette fonction de pénalité en augmentant les valeurs du paramètre α .

Considérons le problème pénalisé :

$$\left[\begin{array}{l} \text{Minimiser } f_{\alpha}(x) \\ \text{sous } x \in \overset{\circ}{S} \end{array} \right.$$

Nous associons à chaque valeur du paramètre α un point central défini par

$$x(\alpha) = \operatorname{argmin}_{x \in \overset{\circ}{S}} f_{\alpha}(x)$$

La courbe $\alpha \in \Re_{++} \rightarrow x(\alpha)$ est la trajectoire centrale pour le problème (PL). Elle a la propriété importante que lorsque α croît, elle converge vers une solution optimale de (PL).

Le comportement polynomial de l'algorithme est obtenu en arrêtant chaque cycle interne, responsable de la minimisation de la fonction de pénalité intérieure, lorsque les itérés sont proches de la trajectoire centrale.

Chaque itération principale met à jour le paramètre de pénalité par un facteur constant et la borne de complexité totale dépend de ce facteur :

Nous obtenons $O(nL)$ itérations pour une constante arbitraire et $O(\sqrt{n}L)$ itérations pour une constante dépendant de \sqrt{n} .

2. L'algorithme de minimisation interne

Nous avons choisi comme direction pour l'algorithme de minimisation interne la direction S-S-D. [cfr chap.IV].

Nous pouvons établir pour la fonction de pénalité intérieure le résultat important suivant:

" Les directions S-S-D et Newton-Raphson sont équivalentes."

Avant de montrer cette équivalence, il est bon de s'intéresser à quelques propriétés concernant la direction Newton-Raphson.

2.1. La direction Newton-Raphson

Nous considérons le problème

$$\left[\begin{array}{l} \text{Minimiser } f(x) \\ \text{sous } A x = b \end{array} \right.$$

où $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est de classe C^2

$A \in \mathbb{R}^{m \times n}$ où $m < n$, de rang m .

Nous appelons direction Newton en un point admissible \bar{x} la solution du sous-problème :

$$(N_{\bar{x}}) \equiv \left[\begin{array}{l} \text{Minimiser } \nabla f(\bar{x})^T d + \frac{1}{2} d^T \nabla^2 f(\bar{x}) d \\ \text{sous } A d = 0 \end{array} \right.$$

Proposition 2.1.1.

Les deux assertions suivantes sont équivalentes :

a. d est solution de $(N_{\bar{x}})$

b. $Ad = 0$ et $P_A [\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d] = 0$

Démonstration

Nous savons, par les conditions nécessaires et suffisantes d'optimalité que

d est solution de (N_x) si et seulement si

$$\exists \lambda \in \mathbb{R}^m \text{ tel que } \nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d + A^T \lambda = 0. \quad (37)$$

Recherchons la valeur de λ :

Par (37), $A^T \lambda = -[\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d]$.

Appliquons la matrice A à chaque membre de l'égalité pour obtenir,

$$A A^T \lambda = -A [\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d].$$

Comme la matrice A est de rang m , la matrice $A A^T$ est inversible.

Ainsi, nous obtenons

$$\lambda = -(A A^T)^{-1} A [\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d]$$

de sorte que

d est solution de (N_x) si et seulement si d vérifie l'équation

$$\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d - A^T (A A^T)^{-1} A [\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d] = 0.$$

Par la mise en évidence de $\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d$, l'équation devient

$$(I - A^T (A A^T)^{-1} A) [\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d] = 0$$

ou encore,

$$P_A [\nabla f(\bar{x}) + \nabla^2 f(\bar{x}) d] = 0$$

Où P_A est la matrice de projection sur le noyau de A . ■

Invariance de la direction Newton par une transformation "scaling"

Soit $D \in \mathbb{R}^{n \times n}$ une matrice diagonale positive inversible. Considérons la transformation de coordonnées :

$$\begin{aligned} D^{-1} : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\rightarrow D^{-1} x = y \end{aligned}$$

Le problème

$$(P) \equiv \begin{cases} \text{Minimiser } f(x) \\ \text{sous } \begin{aligned} A x &= b \\ x &\geq 0 \end{aligned} \end{cases}$$

devient après transformation :

$$(\bar{P}) \equiv \left[\begin{array}{ll} \text{Minimiser} & \bar{f}(y) \\ \text{sous} & \bar{A} y = b \\ & y \geq 0 \end{array} \right] \quad \text{ou} \quad \left[\begin{array}{ll} \text{Minimiser} & f(Dy) \\ \text{sous} & A D y = b \\ & y \geq 0 \end{array} \right]$$

Lemme 2.1.2.

$$P_A [I - D^T P_{AD} D^T] = 0$$

Démonstration

Par définition de la projection sur le noyau de $A D$, nous avons

$$\begin{aligned} D^T P_{AD} D^T &= D^T [I - (A D)^T [A D D^T A^T]^{-1} A D] D^T \\ &= I - D^T D^T A^T [A D D^T A^T]^{-1} A D D^T. \end{aligned}$$

Or $D^T D^T = I$, dès lors,

$$D^T P_{AD} D^T = I - A^T [A D D^T A^T]^{-1} A D D^T.$$

Il suit que

$$P_A [I - D^T P_{AD} D^T] = [I - A^T (A A^T)^{-1} A] A^T [A D D^T A^T]^{-1} A D D^T.$$

En développant cette expression, nous obtenons

$$\begin{aligned} P_A [I - D^T P_{AD} D^T] &= A^T [A D D^T A^T]^{-1} A D D^T - \\ &\quad A^T (A A^T)^{-1} A A^T [A D D^T A^T]^{-1} A D D^T. \end{aligned}$$

Or $(A A^T)^{-1} A A^T = I$, il vient alors que

$$\begin{aligned} P_A [I - D^T P_{AD} D^T] &= A^T [A D D^T A^T]^{-1} A D D^T - \\ &\quad A^T [A D D^T A^T]^{-1} A D D^T = 0. \quad \blacksquare \end{aligned}$$

Proposition 2.1.3.

La direction de Newton, d , est invariante par une transformation "scaling".

Démonstration

Soit $\bar{y} \geq 0$, un point admissible du problème (\bar{P}) .

Autrement dit, $A D \bar{y} = b$.

La direction de Newton dans l'espace transformé en $\bar{y} = D^{-1} x$ est solution de

$$(N_{\bar{y}}) \equiv \begin{cases} \text{Minimiser } \nabla \bar{f}(\bar{y})^T u + \frac{1}{2} u^T \nabla^2 \bar{f}(\bar{y}) u \\ \text{sous} \quad A D u = 0 \end{cases}$$

Développons la fonction objectif de $(N_{\bar{y}})$ en considérant :

$$\begin{aligned} \text{i) } \nabla \bar{f}(\bar{y}) &= \nabla f \circ D(\bar{y}) \\ &= D^T \nabla f(D \bar{y}) \\ &= D^T \nabla f(x) \end{aligned}$$

$$\begin{aligned} \text{ii) } \nabla^2 \bar{f}(\bar{y}) &= \nabla^2 f \circ D(\bar{y}) \\ &= \nabla (D^T \nabla f(D \bar{y})) \\ &= D^T \nabla^2 f(D \bar{y}) D \\ &= D^T \nabla^2 f(x) D \end{aligned}$$

Le problème $(N_{\bar{y}})$ peut alors s'écrire

$$\begin{cases} \text{Minimiser } D^T \nabla f(x)^T u + \frac{1}{2} u^T D^T \nabla^2 f(x) D u \\ \text{sous} \quad A D u = 0 \end{cases}$$

Soit u la solution de (N_y) . Montrer que la direction Newton, d , est invariante par scaling revient à montrer que $d = D u$.

Par la proposition 2.1.1, nous avons l'équivalence suivante

u est solution de (N_y)

\Leftrightarrow

$$A D u = 0 \text{ et } P_{AD} [D^T \nabla f(\bar{x}) + D^T \nabla^2 f(\bar{x}) D u] = 0$$

ou encore

$$A D u = 0 \text{ et } P_A D^T P_{AD} D^T [D^T \nabla f(\bar{x}) + D^T \nabla^2 f(\bar{x}) D u] = 0.$$

Par le lemme 2.1.2., les conditions d'optimalité deviennent

$$A D u = 0 \text{ et } P_A [D^T \nabla f(\bar{x}) + D^T \nabla^2 f(\bar{x}) D u] = 0$$

Autrement dit, $D u$ est solution de (N_x) et donc $D u = d$. ■

2.2. Coïncidence des directions S-S-D et Newton-Raphson

Proposition 2.2.1.

Considérons la fonction de pénalité intérieure $f_\alpha(x)$ pour $\alpha \geq 0$ fixé.
La direction Newton en \bar{x} coïncide avec la direction S.S.D,
donnée par

$$i) h_N(e, \alpha) = - P_A \nabla f_\alpha(e) = - \alpha c_p + e_p$$

$$\begin{aligned} ii) h_N(\bar{x}, \alpha) &= - X P_{AX} X \nabla f_\alpha(\bar{x}) \\ &= - X P_A^-(\alpha \bar{c} - e) \end{aligned}$$

Démonstration

La direction de Newton en \bar{x} telle que

$$\bar{x} > 0 \text{ et } A \bar{x} = b$$

est donnée par la résolution du problème

$$(N_x) = \begin{cases} \text{Minimiser } \nabla f_\alpha(\bar{x})^T d + \frac{1}{2} d^T \nabla^2 f_\alpha(\bar{x}) d \\ \text{sous} \quad A d = 0 \end{cases}$$

i) Supposons que $\bar{x} = e$; ceci est possible si $A e = b$.

Le problème (N_e) s'écrit alors, puisque $\nabla^2 f_\alpha(e) = I$,

$$\begin{cases} \text{Minimiser } \nabla f_\alpha(e)^T d + \frac{1}{2} d^T d \\ \text{sous} \quad A d = 0 \end{cases}$$

La proposition 2.1.1 nous donnent des conditions nécessaires et suffisantes d'optimalité:

d est solution de (N_e)

\Leftrightarrow

$$A d = 0 \text{ et } P_A [\nabla f_\alpha(e) + d] = 0.$$

L'opérateur projection est linéaire; nous avons donc

$$A d = 0 \text{ et } -P_A \nabla f_\alpha(e) = P_A d.$$

Or $P_A d = d$ car $d \in \text{Ker } A$. Nous obtenons ainsi

$$d = -P_A \nabla f_\alpha(e).$$

ii) Pour calculer la solution de (N_x) , nous effectuons une transformation " scaling " :

$$\begin{aligned} X^{-1} : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\rightarrow X^{-1} x = y \end{aligned}$$

où $X = \text{diag}(\bar{x})$

Le point x est, de cette façon, envoyé sur le point e par " scaling ".

La direction de Newton en e dans l'espace transformé, notée $\bar{h}_N(e, \alpha)$, est solution du problème

$$\left[\begin{array}{l} \text{Minimiser } \nabla \bar{f}_\alpha(e) u + \frac{1}{2} u^T u \\ \text{sous} \quad A X u = 0 \end{array} \right.$$

Dès lors, par (i),

$$\begin{aligned} \bar{h}_N(e, \alpha) &= -P_{AX} \nabla \bar{f}_\alpha(e) \\ &= -P_{AX} X \nabla f_\alpha(\bar{x}). \end{aligned}$$

Comme la direction Newton est invariante par "scaling" (proposition 2.1.3),

$$h_N(\bar{x}, \alpha) = X \bar{h}_N(e, \alpha).$$

C'est-à-dire

$$h_N(\bar{x}, \alpha) = -X P_{AX} X \nabla f_\alpha(\bar{x}). \quad \blacksquare$$

3. Le critère de proximité

Définition 3.1.

Soit $\alpha > 0$ et $x \in \overset{\circ}{S}$,
 La mesure de proximité d'un point x par rapport au point central $x(\alpha)$ est donnée par

$$\delta(x, \alpha) = \|X^{-1} h_N(x, \alpha)\|$$

La mesure de proximité décrit la longueur du pas Newton obtenu après un "scaling" amenant le point courant sur le vecteur unité.

Le critère de proximité peut s'énoncer alors:

"Le point x est considéré proche du point central $x(\alpha)$ si $\delta(x, \alpha) \leq 0,1$."

Proposition 3.2.

Si $\delta(x, \alpha) < 1$, le pas Newton nous dirige vers un nouveau point intérieur.

Démonstration

Soit x , un point courant.

Après un pas Newton, le nouveau point est $x + h_N(x, \alpha)$.

Considérons la transformation " scaling " X^{-1} où $X = \text{diag}(x_i)$.

Nous avons

$$\begin{aligned}\delta(x, \alpha) &= \|X^{-1} h_N(x, \alpha)\| \\ &= \|\bar{h}_N(e, \alpha)\| \\ &< 1.\end{aligned}$$

Ainsi,

$$e + \bar{h}_N(e, \alpha) > 0.$$

Il suit que

$X e + X \bar{h}_N(e, \alpha)$
reste strictement positif, puisque X est une matrice diagonale positive. Dès lors,
 $x + h_N(x, \alpha) > 0$. ■

Les principales propriétés associées au concept de proximité sont résumées dans la proposition suivante :

Proposition 3.3.

Considérons un point $x \in \mathring{S}$ et $\delta(x, \alpha)$, sa mesure de proximité par rapport au point central $x(\alpha)$.

Nous pouvons établir les propriétés suivantes :

$$1) \text{ Si } \delta(x, \alpha) \leq 0,1 \text{ alors } \|X^{-1}(x - x(\alpha))\| \leq 0,115$$

$$2) \text{ Si } \delta(x, \alpha) \leq 0,1 \text{ alors } f_{\alpha}(x) - f_{\alpha}(x(\alpha)) \leq 0,012$$

$$3) \text{ Si } \delta(x, \alpha) \leq 0,1 \text{ alors } |c^T x - c^T x(\alpha)| \leq 0,2 \frac{\sqrt{n}}{\alpha}$$

$$4) \text{ Si } \delta(x, \alpha) \geq 0,1 \text{ alors } f_{\alpha}(y) \leq f_{\alpha}(x) - 0,004,$$

$$\text{où } y = x + \frac{0,1}{\delta(x, \alpha)} h_N(x, \alpha)$$

Démonstration

Soit $x > 0$, tel que $Ax = b$.

Considérons la transformation " scaling " :

$$X^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$x \rightarrow X^{-1}x = y$$

$$\text{où } X = \text{diag}(x_1 \dots x_n)$$

Nous pouvons montrer que :

$$i) \ x(\alpha) \text{ solution de (P)} \equiv \begin{cases} \text{Minimiser } f_{\alpha}(x) \\ \text{sous } Ax = b \end{cases}$$

\Leftrightarrow

$$\bar{x}(\alpha) \text{ solution de } (\bar{P}) \equiv \begin{cases} \text{Minimiser } \bar{f}_{\alpha}(y) \\ \text{sous } \bar{A}y = b \end{cases}$$

$$\text{ou } \bar{P} \equiv \begin{cases} \text{Minimiser } f_{\alpha}(Xy) \\ \text{sous } AXy = b \end{cases}$$

$$\text{où } \bar{x}(\alpha) = X^{-1}x(\alpha).$$

En effet, la fonction de pénalité intérieure est une fonction strictement convexe. Les conditions nécessaires et suffisantes d'optimalité nous donne alors les équivalences suivantes :

$$\begin{aligned}
 & \bar{x}(\alpha) \text{ est solution de (P)} \\
 \Leftrightarrow & \\
 & \exists \lambda \in \mathbb{R}^m \quad A \bar{x}(\alpha) = b \text{ et } \nabla f_\alpha(\bar{x}(\alpha)) + A^T \lambda = 0 \\
 \Leftrightarrow & \\
 & \exists \lambda \in \mathbb{R}^m \quad A X \bar{x}(\alpha) = b \text{ et } X \nabla f_\alpha(X \bar{x}(\alpha)) + X A^T \lambda = 0 \\
 \Leftrightarrow & \\
 & \bar{x}(\alpha) \text{ est solution de } (\bar{P}).
 \end{aligned}$$

$$\text{ii) } \delta(\bar{x}, \alpha) = \bar{\delta}(e, \alpha)$$

En effet,

$$\delta(\bar{x}, \alpha) = \|X^{-1} h_N(\bar{x}, \alpha)\| \quad \text{et} \quad \bar{\delta}(e, \alpha) = \|\bar{h}_N(e, \alpha)\|.$$

Or $\bar{h}_N(e, \alpha) = X^{-1} h_N(\bar{x}, \alpha)$. Il suit que

$$\bar{\delta}(e, \alpha) = \delta(\bar{x}, \alpha).$$

En vertu des résultats obtenus aux points i) et ii), nous pouvons réécrire les différentes propriétés du lemme sous la forme :

Si $\bar{\delta}(e, \alpha) \leq 0,1$,

$$1) \|e - \bar{x}(\alpha)\| \leq 0,115$$

$$2) \bar{f}_\alpha(e) - \bar{f}_\alpha(\bar{x}(\alpha)) \leq 0,012$$

$$3) \left| \bar{c}^T e - \bar{c}^T \bar{x}(\alpha) \right| \leq 0,2 \frac{\sqrt{n}}{\alpha}$$

Si $\bar{\delta}(e, \alpha) \geq 0,1$,

$$4) \bar{f}_\alpha(\bar{y}) \leq \bar{f}_\alpha(e) - 0,004$$

$$\text{où } \bar{y} = e + \frac{0,1}{\bar{\delta}(e, \alpha)} \bar{h}_N(e, \alpha).$$

1) Considérons la direction normalisée définie comme suit

$$h = \frac{\bar{x}(\alpha) - e}{\|\bar{x}(\alpha) - e\|}$$

- Nous nous proposons, dans un premier temps, d'étudier la fonction F définie par

$$F(\lambda) = \bar{f}_\alpha(e + \lambda h) \text{ où } \lambda \in [0, 1).$$

a) F est strictement convexe

Par définition de la convexité d'une fonction, nous devons montrer que

$$\forall \lambda_1, \lambda_2 \in [0, 1), \forall \mu \in (0, 1) :$$

$$F(\mu \lambda_1 + (1 - \mu) \lambda_2) < \mu F(\lambda_1) + (1 - \mu) F(\lambda_2).$$

Autrement dit, que

$$\bar{f}_\alpha(e + (\mu \lambda_1 + (1 - \mu) \lambda_2)h) < \mu \bar{f}_\alpha(e + \lambda_1 h) + (1 - \mu) \bar{f}_\alpha(e + \lambda_2 h)$$

ou encore de façon équivalente, que

$$f_\alpha[X(e + (\mu \lambda_1 + (1 - \mu) \lambda_2)h)] < \mu f_\alpha(X(e + \lambda_1 h)) + (1 - \mu) f_\alpha(X(e + \lambda_2 h)).$$

Par définition de la fonction de pénalité intérieure, la thèse devient

$$\begin{aligned} \alpha c^T (X(e + (\mu \lambda_1 + (1 - \mu) \lambda_2)h)) + p[X(e + (\mu \lambda_1 + (1 - \mu) \lambda_2)h)] \\ < \mu \alpha c^T X(e + \lambda_1 h) + \mu p(X(e + \lambda_1 h)) \\ + (1 - \mu) \alpha c^T X(e + \lambda_2 h) + (1 - \mu) p(X(e + \lambda_2 h)) \end{aligned}$$

En développant les membres de cette inégalité et après quelques simplifications, nous obtenons les équivalences suivantes :

$$\begin{aligned} p[X(e + (\mu \lambda_1 + (1 - \mu) \lambda_2)h)] &< \mu p(X(e + \lambda_1 h)) + (1 - \mu) p(X(e + \lambda_2 h)) \\ \Leftrightarrow p(e + (\mu \lambda_1 + (1 - \mu) \lambda_2)h) &< \mu p(e + \lambda_1 h) + (1 - \mu) p(e + \lambda_2 h) \\ \Leftrightarrow p(\mu(e + \lambda_1 h) + (1 - \mu)(e + \lambda_2 h)) &< \mu p(e + \lambda_1 h) + (1 - \mu) p(e + \lambda_2 h) \quad (38) \end{aligned}$$

Montrer que F est strictement convexe revient donc à montrer que l'inégalité (38) est vérifiée pour tout $\lambda_1, \lambda_2 \in [0, 1)$ et $\mu \in (0, 1)$.

Ceci est immédiat puisque la fonction barrière logarithmique est strictement convexe [cfr chap.III]

b) $\bar{\lambda} = \|\bar{x}(\alpha) - e\|$ est l'unique minimum de F

En effet, nous avons d'une part

$$\begin{aligned} \frac{d}{d\lambda} F(\bar{\lambda}) &= \nabla (\bar{f}_\alpha (e + \bar{\lambda} h)) \\ &= h^T \nabla \bar{f}_\alpha (e + \bar{\lambda} h) \end{aligned}$$

$$= \frac{(\bar{x}(\alpha) - e)^T}{\|\bar{x}(\alpha) - e\|} \nabla \bar{f}_\alpha (\bar{x}(\alpha))$$

Or $\bar{x}(\alpha)$ est solution de (\bar{P}) . Dès lors,

$$P_{A \times} \nabla \bar{f}_\alpha (\bar{x}(\alpha)) = 0$$

ou encore

$\nabla \bar{f}_\alpha (\bar{x}(\alpha))$ est orthogonal au noyau de $A \times$.

Par ailleurs, puisque $A \times (\bar{x}(\alpha) - e) = A \times \bar{x}(\alpha) - A \times e = 0$, nous pouvons dire que

$$(\bar{x}(\alpha) - e) \in \text{Ker}(A \times)$$

et donc

$$(\bar{x}(\alpha) - e)^T \nabla \bar{f}_\alpha (\bar{x}(\alpha)) = 0.$$

Autrement dit

$$\frac{d}{d\lambda} F(\bar{\lambda}) = 0.$$

D'autre part, la fonction F est strictement convexe;

il suit que $\bar{\lambda}$ est le minimum unique de F .

c) $\frac{d}{d\lambda} F(\lambda) = h^T \nabla \bar{f}_\alpha (e + \lambda h)$ est une fonction croissante monotone.

En effet, nous pouvons montrer que la dérivée de cette fonction, c'est-à-dire la dérivée seconde de $F(\lambda)$ est positive.

$$\begin{aligned}
\frac{d^2}{d\lambda^2} \bar{f}_\alpha(e + \lambda h) &= \nabla (h^T \nabla \bar{f}_\alpha(e + \lambda h)) \\
&= h^T \nabla^2 \bar{f}_\alpha(e + \lambda h) h \\
&= h^T X^T \nabla^2 f_\alpha(X(e + \lambda h)) X h.
\end{aligned}$$

Puisque la matrice $\nabla^2 f_\alpha(x)$ est définie positive pour $x > 0$, nous avons bien

$$\frac{d^2}{d\lambda} F(\lambda) \geq 0.$$

- Après ces quelques considérations sur la fonction F ,
montrons que $\bar{\lambda} < 0,115$.

Examinons pour cela le signe de $\frac{d}{d\lambda} \bar{f}_\alpha(e + \lambda h)$ pour un λ suffisamment petit.

$$\begin{aligned}
\frac{d}{d\lambda} \bar{f}_\alpha(e + \lambda h) &= h^T \nabla \bar{f}_\alpha(e + \lambda h) \\
&= h^T (\nabla \bar{f}_\alpha(e) + \lambda h + o(\lambda h)) \\
&\quad [\text{cfr chap III: prop 3.1.1 + note}] \\
&= h^T \nabla \bar{f}_\alpha(e) + \lambda + h^T o(\lambda h)
\end{aligned}$$

Or, nous avons les relations suivantes :

$$|h^T o(\lambda h)| \leq \|h^T\| \|o(\lambda h)\| = \|o(\lambda h)\|$$

et

$$\|o(\lambda h)\| \leq \frac{\lambda^2}{1 - \lambda}.$$

Dès lors,

$$h^T o(\lambda h) \geq \frac{-\lambda^2}{1 - \lambda}.$$

De ce résultat, il suit que

$$\frac{d}{d\lambda} \bar{f}_\alpha(e + \lambda h) \geq h^T \nabla \bar{f}_\alpha(e) + \lambda - \frac{\lambda^2}{1 - \lambda}.$$

Minorons encore cette inégalité en sachant que

$$|h^T \nabla \bar{f}_\alpha(e)| = |h^T P \nabla \bar{f}_\alpha(e)| \leq \|P_{\bar{A}} \nabla \bar{f}_\alpha(e)\| \leq 0,1$$

car $h \in \text{Ker}(\bar{A})$ et P est une projection orthogonale sur \bar{A} et par conséquent,

$$h^T \nabla \bar{f}_\alpha(e) \geq -0,1 \quad (39).$$

Nous obtenons donc finalement

$$\frac{d}{d\lambda} \bar{f}_\alpha(e + \lambda h) \geq -0,1 + \lambda - \frac{\lambda^2}{1-\lambda}.$$

Prenons $\lambda = 0,115$, nous avons alors

$$\frac{d}{d\lambda} \bar{f}_\alpha(e + 0,115 h) > 0$$

et puisque la dérivée, $\frac{d}{d\lambda} \bar{f}_\alpha(e + \lambda h)$, est une fonction croissante, elle croise 0 en $\bar{\lambda} < 0,115$. Autrement dit,

$$\|\bar{x}(\alpha) - e\| < 0,115.$$

2) En vertu de la proposition 3.1.1. [cfr. chap.III] et en considérant la note qui la suit, nous avons la relation suivante

$$\bar{f}_\alpha(e + \lambda h) \geq \bar{f}_\alpha(e) + \lambda \nabla \bar{f}_\alpha(e)^T h.$$

Or $\nabla \bar{f}_\alpha(e)^T h \geq -0,1$ par (39) et prenons $\lambda \leq 0,115$.
Nous obtenons ainsi que

$$\begin{aligned} \bar{f}_\alpha(e + \lambda h) - \bar{f}_\alpha(e) &\geq (-0,1) \cdot (0,115) \\ &= -0,0115 \end{aligned}$$

ou

$$\bar{f}_\alpha(e) - \bar{f}_\alpha(e + \lambda h) \leq 0,0115.$$

3) Définissons $d : d = e - \bar{x}(\alpha)$

Par le point 1) de la proposition, nous savons que $\|d\| \leq 0,115$.

Par conséquent, comme $d \in \text{Ker}(\bar{A})$, nous avons

$$\left| d^T \nabla \bar{f}_\alpha(e) \right| = \left| d^T P \nabla \bar{f}_\alpha(e) \right|$$

et par l'inégalité de Cauchy-Schwartz,

$$\begin{aligned} \left| d^T \nabla \bar{f}_\alpha(e) \right| &\leq \|d\| \left\| P \nabla \bar{f}_\alpha(e) \right\| \\ &\leq 0,115 \cdot 0,1 = 0,0115 \end{aligned}$$

Or $\nabla \bar{f}_\alpha(e) = \alpha \bar{c} - e$, d'où

$$\left| \alpha \bar{c}^T d - e^T d \right| \leq 0,0115$$

c'est-à-dire

$$-0,0115 \leq \alpha \bar{c}^T d - e^T d \leq 0,0115.$$

Comme $|e^T d| \leq \|e\| \|d\| \leq \sqrt{n} \cdot 0,115$, nous obtenons

$$-0,0115 - \sqrt{n} \cdot 0,115 \leq \alpha \bar{c}^T d \leq 0,115 \sqrt{n} + 0,0115$$

et finalement

$$\left| \alpha \bar{c}^T d \right| \leq (0,115 + 0,0115) \sqrt{n} \leq 0,2 \sqrt{n}.$$

Autrement dit, il vient que

$$\left| \bar{c}^T e - \bar{c}^T \bar{x}(\alpha) \right| \leq 0,2 \frac{\sqrt{n}}{\alpha}.$$

4) Soit $\bar{y} = e + 0,1 h$

$$\text{où } h = \frac{-P \nabla \bar{f}_\alpha(e)}{\|P \nabla \bar{f}_\alpha(e)\|} = \frac{\bar{h}_N(e, \alpha)}{\bar{\delta}(e, \alpha)}.$$

- Nous pouvons vérifier que \bar{y} est admissible:

en effet, la norme de h vaut 1,
chacune de ses composantes est alors comprise entre -1 et +1,
et donc

$$\bar{y} \geq 0$$

De plus,

$$\begin{aligned} A X \bar{y} &= A X (e + 0,1 h) \\ &= A \bar{x} + 0,1 A X h \\ &= b + 0 \quad (h \in \text{Ker}(AX)) \\ &= b \end{aligned}$$

- Evaluons $\bar{f}_\alpha(\bar{y})$.

Nous obtenons les inégalités suivantes

$$\bar{f}_\alpha(\bar{y}) \leq \bar{f}_\alpha(e) + \lambda \nabla \bar{f}_\alpha(e)^T h + \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|} \quad \text{où } \lambda = 0,1$$

[cfr chapitre III prop. 3.1.1. + note].

$$\leq \bar{f}_\alpha(e) + 0,1 \nabla \bar{f}_\alpha(e)^T h + \frac{0,01}{1,8}.$$

$$\text{Or } h = \frac{-P \nabla \bar{f}_\alpha(e)}{\|P \nabla \bar{f}_\alpha(e)\|} \text{ et}$$

$$\nabla \bar{f}_\alpha(e)^T P \nabla \bar{f}_\alpha(e) = \nabla \bar{f}_\alpha(e)^T P^T P \nabla \bar{f}_\alpha(e) = \|P \nabla \bar{f}_\alpha(e)\|^2$$

(P est la matrice de projection orthogonale et $PP = P$).
Nous avons ainsi que

$$\bar{f}_\alpha(\bar{y}) \leq \bar{f}_\alpha(e) - 0,1 \|P \nabla \bar{f}_\alpha(e)\| + \frac{0,01}{1,8}.$$

Puisque $\|P \nabla \bar{f}_\alpha(e)\| \geq 0,1$ par hypothèse, il suit immédiatement que

$$\bar{f}_\alpha(\bar{y}) \leq \bar{f}_\alpha(e) - 0,004 \quad \blacksquare$$

4. L'algorithme pouvant être mis en œuvre

L'algorithme est précisément la méthode barrière avec un critère de proximité pour rompre les itérations internes après chaque mise à jour du paramètre de pénalité.

L'algorithme se divise en deux parties :

1) l'algorithme principal

A l'itération k, il met à jour le paramètre de pénalité α_k par un facteur constant $(1 + \theta)$ avec

$$\theta = \frac{\gamma}{\sqrt{n}}.$$

$$\text{Ainsi } \alpha_{k+1} = (1 + \theta) \alpha_k$$

2) Pour un α_k :
l'algorithme interne

Il applique la méthode Newton pour déterminer une approximation du minimum de la fonction pénalisée f_{α_k}

Son critère d'arrêt est un critère de proximité du point courant par rapport au point central associé au paramètre α_k .

4.1. Le critère d'arrêt l'algorithme principal

L'algorithme principal s'arrête lorsqu'un point x^k est obtenu et vérifie

$$\delta(x^k, \alpha_k) \leq 0,1 \text{ et } \alpha_k \geq 1,2 n 2^L$$

En effet, nous savons par la proposition 4.2.1.2 [cfr. chap. III] que le saut de dualité s'exprime comme

$$c^T(x(\alpha)) - b^T y(\alpha) = \frac{n}{\alpha}.$$

(nous notons $\alpha \equiv \alpha_k$)

Or $\hat{v} \geq b^T y(\alpha)$, par conséquent

$$c^T(x(\alpha)) - \hat{v} \leq \frac{n}{\alpha}. \quad (40)$$

Par ailleurs $\delta(x^k, \alpha_k) \leq 0,1$.

En se référant à la proposition 3.3, nous avons la relation

$$|c^T x^k - c^T x(\alpha)| \leq 0,2 \frac{\sqrt{n}}{\alpha} \quad (41)$$

En sommant les deux inégalités (40) et (41), nous obtenons

$$\begin{aligned} c^T x^k - \hat{v} &\leq \frac{n + 0,2 \sqrt{n}}{\alpha} \\ &\leq \frac{1,2 n}{\alpha}. \end{aligned}$$

Or $\alpha \geq 1,2 n 2^L$, dès lors

$$c^T x^k - \hat{v} \leq 2^{-L}.$$

Nous savons qu'à ce moment, il suffit d'appliquer une procédure de "purification" pour obtenir une solution optimale.

4.2. L'algorithme

Nous supposons que l'algorithme commence avec un paramètre

$$\alpha_0 \geq 2^{-L}$$

et un point x^0 tel que $\delta(x^0, \alpha_0) \leq 0,1$

Algorithme

Soient, donnés: $\theta = \frac{\gamma}{\sqrt{n}}$ ($\gamma > 0$),

$x^0 \in S^0$ et $\alpha_0 \geq 0$ tels que $\delta(x^0, \alpha_0) \leq 0,1$.

- $k = 0$

"algorithme principal"

- Répéter jusqu'à ce que $\alpha_k > 1,2 n 2^L$

- $\bar{\alpha} = (1 + \theta) \alpha_k$

- $j = 0$

- $y^0 = x^k$

- Répéter jusqu'à ce que $\delta(y^j, \bar{\alpha}) \leq 0,1$

"algorithme interne"

- Calcul de la direction: $h = h_N(y^j, \bar{\alpha})$

- Recherche linéaire

$$\bar{\lambda} = \operatorname{argmin} \{f_{\alpha}(y^j + \lambda h) / \lambda \geq 0\}$$

$$y^{j+1} = y^j + \bar{\lambda} h$$

- $j = j + 1$

- $x^{k+1} = y^j$

- $\alpha_{k+1} = \bar{\alpha}$

- $k = k + 1$

5. Etude de la complexité de l'algorithme

Nous étudions la complexité de l'algorithme en considérant la complexité de chacune de ses parties : l'algorithme principal et l'algorithme interne.

Proposition 5.1.

L'algorithme s'arrête en au plus $\frac{1 + \theta}{\theta} O(L)$ itérations principales.

Démonstration

Plaçons - nous à l'itération k , le paramètre α_k satisfait les relations

$$\alpha_k \leq 1,2 n 2^L \text{ et } \alpha_k = (1 + \theta) \alpha_{k-1}.$$

Nous pouvons encore exprimer α_k comme $\alpha_k = (1 + \theta)^k \alpha_0$.

Nous avons donc la relation suivante

$$(1 + \theta)^k \alpha_0 \leq 1,2 n 2^L.$$

En prenant le logarithme de base 2 de chaque membre de l'inégalité, nous avons

$$k \log_2 (1 + \theta) \leq L + \log_2 (1,2 n) - \log_2 \alpha_0. \quad (42)$$

Or la fonction logarithme est une fonction concave, autrement dit,

$$\log (1 + \theta) \geq \theta \frac{d}{d\theta} \log (1 + \theta) = \frac{\theta}{1 + \theta}. \quad (43)$$

Il suit de (42) et de (43) que

$$k \frac{\theta}{1 + \theta} \leq L + \log_2 (1,2 n) - \log_2 \alpha_0.$$

La condition initiale sur le paramètre $\alpha_0 > 2^{-L}$

nous permet de conclure que $k \leq \frac{1 + \theta}{\theta} O(L)$. ■

Proposition 5.2.

Considérons une itération de l'algorithme principal avec

$$\bar{\alpha} = (1 + \theta) \alpha_k \text{ et soit } \Delta = 0,004.$$

Le nombre J de pas internes satisfait :

$$J \Delta \leq \frac{\theta^2}{1 + \theta} n + \theta \sqrt{n} + 1.$$

Démonstration

Soit k une itération de l'algorithme principal.

Notons $\alpha \equiv \alpha_k$ et $\bar{\alpha} \equiv (1 + \theta) \alpha$.

L'algorithme interne démarre avec y^0 et s'arrête lorsqu'il a obtenu y^J

tel que $\delta(y^J, \bar{\alpha}) \leq 0,1$.

Pour $i = 0, \dots, J-1$, nous avons $\delta(y^i) \geq 0,1$.

Nous appliquons alors la proposition 3.3 (4), pour obtenir successivement

$$f_{\bar{\alpha}}(y^1) \leq f_{\bar{\alpha}}(y^0) - \Delta,$$

$$f_{\bar{\alpha}}(y^2) \leq f_{\bar{\alpha}}(y^1) - \Delta \leq f_{\bar{\alpha}}(y^0) - 2\Delta, \text{ etc.}$$

Il suit facilement que

$$f_{\bar{\alpha}}(y^J) \leq f_{\bar{\alpha}}(y^0) - J\Delta. \quad (44)$$

En outre, comme $\bar{\alpha} = (1 + \theta) \alpha$, nous avons

$$\begin{aligned} f_{\bar{\alpha}}(y) &= (1 + \theta) \alpha c^T y + p(y) \\ &= \alpha c^T y + p(y) + \theta \alpha c^T y \\ &= f_{\alpha}(y) + \theta \alpha c^T y \quad \forall y \in \mathring{S}. \end{aligned}$$

Ainsi l'inégalité (44) s'exprime maintenant comme

$$f_{\alpha}(y^J) + \theta \alpha c^T y^J \leq f_{\alpha}(y^0) + \theta \alpha c^T y^0 - J\Delta$$

ou encore

$$J\Delta \leq \theta \alpha c^T y^0 - \theta \alpha c^T y^J + f_{\alpha}(y^0) - f_{\alpha}(y^J).$$

Or $\delta(y^o, \alpha) \leq 0,1$ car $y^o = x^k$ est proche du point central $x(\alpha)$.

De plus, $f_\alpha(y^j) \geq f_\alpha(x(\alpha))$.

Nous obtenons alors par la proposition 3.3 (2),

$$f_\alpha(y^o) - f_\alpha(y^j) \leq 0,012 < 1.$$

Ainsi,

$$J \Delta \leq \theta \alpha c^T y^o - \theta \alpha c^T y^j + 1$$

$$\text{ou même, } \frac{J \Delta - 1}{\theta} \leq \alpha c^T y^o - \alpha c^T y^j.$$

Poursuivons la majoration de $J \Delta$:

Comme $\delta(y^o, \alpha) \leq 0,1$ et $\delta(y^j, \bar{\alpha}) \leq 0,1$, en vertu de la proposition 3.3 (3) nous assurant les relations

$$c^T y^o \leq c^T x(\alpha) + \frac{0,2 \sqrt{n}}{\alpha}$$

$$\text{et } c^T y^j \geq c^T x(\bar{\alpha}) - \frac{0,2 \sqrt{n}}{\bar{\alpha}}$$

nous pouvons affirmer que

$$\frac{J \Delta - 1}{\theta} \leq \alpha (c^T x(\alpha) + \frac{0,2 \sqrt{n}}{\alpha}) - \alpha (c^T x(\bar{\alpha}) - \frac{0,2 \sqrt{n}}{\bar{\alpha}}).$$

Après simplification, il vient que,

$$\frac{J \Delta - 1}{\theta} \leq \alpha c^T x(\alpha) - \alpha c^T x(\bar{\alpha}) + (1 + \frac{1}{1 + \theta}) 0,2 \sqrt{n} \quad . \quad (45)$$

Considérons maintenant les coûts duaux associés à α et $\bar{\alpha}$,

$$v(\alpha) = c^T x(\alpha) - \frac{n}{\alpha} \quad (46)$$

$$v(\bar{\alpha}) = c^T x(\bar{\alpha}) - \frac{n}{\bar{\alpha}}.$$

Ceux-ci garantissent que

$$\alpha (v(\bar{\alpha}) - v(\alpha)) \geq 0 \quad . \quad (47)$$

Nous en déduisons, en additionnant (45) et (47), que

$$\frac{J \Delta - 1}{\theta} \leq \alpha (c^T x(\alpha) - v(\alpha)) - \alpha (c^T x(\bar{\alpha}) - v(\bar{\alpha})) + \left(1 + \frac{1}{1 + \theta}\right) 0,2 \sqrt{n}.$$

Remarquons encore que $\left(1 + \frac{1}{1 + \theta}\right) 0,2 \leq 1$,

$$\text{d'où } \frac{J \Delta - 1}{\theta} \leq \alpha (c^T x(\alpha) - v(\alpha)) - \alpha (c^T x(\bar{\alpha}) - v(\bar{\alpha})) + \sqrt{n}$$

Finalement, nous pouvons utiliser (46) pour obtenir

$$\frac{J \Delta - 1}{\theta} \leq n - \frac{1}{1 + \theta} n + \sqrt{n}$$

ou de façon équivalente,

$$J \Delta - 1 \leq \theta \sqrt{n} + \frac{\theta^2}{1 + \theta} n. \quad \blacksquare$$

CONCLUSION

Le nombre total K d'itérations de l'algorithme complet est donné par le produit des bornes des propositions 5.1 et 5.2, c'est-à-dire

$$K \Delta \leq [n \theta + (1 + \theta) \sqrt{n} + \frac{1 + \theta}{\theta}] O(L)$$

En substituant la valeur de $\theta = \frac{\gamma}{\sqrt{n}}$, nous avons,

$$K \Delta \leq [\gamma \sqrt{n} + \gamma + \sqrt{n} + \frac{\sqrt{n} + \gamma}{\gamma}] O(L)$$

Pour $\gamma = O(1)$, $\gamma > 0,05$, le nombre total de pas Newton est de l'ordre

de $O(\gamma \sqrt{n} L)$.

Nous devons considérer deux cas :

1) Si γ est une constante arbitraire, la borne est $O(\sqrt{n} L)$.

2) Si $\gamma = \theta \sqrt{n}$ où θ est arbitraire, $\theta = O(1)$, la borne de complexité est alors $O(n L)$ itérations.

Chapitre VI

La méthode de la fonction potentielle

1. Préliminaires

L'idée générale de la méthode de la fonction potentielle est très semblable de celle de la méthode de la fonction de pénalité intérieure:

Elle génère une suite de sous-problèmes dont la résolution se fait de manière approximative en utilisant l'algorithme S-S-D.
La suite des points optimaux à chaque sous-problème converge vers la solution du problème initial (PL).

Par ailleurs, cette méthode diffère par sa fonction auxiliaire qui est maintenant la fonction potentielle

$$f_q(x) = q \log(c^T x - v) + p(x)$$

où $x \in \overset{\circ}{S}$.

De plus, le critère de proximité par rapport à la trajectoire centrale est utilisé pour la mise à jour du paramètre v et en même temps permet de réduire le saut de dualité.
L'analyse de complexité est aussi très différente.

Nous présentons, dans ce chapitre, deux algorithmes :

Le premier utilise un critère de proximité très lâche et une mise à jour très flexible; il a une complexité de $O(nL)$ itérations. Le second est plus strict et a une borne de complexité de $O(\sqrt{n} L)$ itérations.

2. La direction S-S-D

Nous avons vu que pour la fonction de pénalité intérieure, la direction S-S-D coïncidait avec la direction Newton-Raphson. Il n'en est pas de même pour la fonction potentielle :

Considérons l'approximation quadratique de la fonction auxiliaire,

$$f_p(x + h) \cong f_p(x) + \nabla f_p(x)^T h + \frac{1}{2} h^T E h$$

La méthode Newton minimise à chaque itération le modèle quadratique de la fonction avec $E = \nabla^2 f_p(x)$. Par contre, il a été observé que l'algorithme S-S-D utilise $E = \nabla^2 p(x)$.

Nous perdons ainsi la contribution des dérivées secondes du premier terme de la fonction auxiliaire. Cette contribution est nulle pour la fonction de pénalité intérieure; c'est pourquoi les méthodes coïncident. Par contre, pour la fonction potentielle, cette contribution est négative. Elle détruit le caractère défini positif de la matrice Hessienne et est donc ignorée.

3. Le critère de proximité et la mise à jour du paramètre v

Selon la définition de la mesure de proximité d'un point, x , par rapport au point central $x(v)$, nous pouvons énoncer le critère de proximité pour la fonction potentielle comme suit :

$$\|\bar{h}(x, v)\| < \varepsilon \quad \text{où } \varepsilon \in (0, 1]$$

c'est-à-dire,

$$\|P \nabla \bar{f}_q(x, v)\| < \varepsilon \quad \text{où } \varepsilon \in (0, 1].$$

Supposons que le point e est le point central associé au paramètre v dans l'espace transformé par un " scaling ".

Le lemme 3.1. nous fournit un saut de dualité pour des points proches de la trajectoire centrale et sera très important dans la procédure de mise à jour du paramètre v .

Lemme 3.1.

Soient $q > n$ et $v_1 < \hat{v}$.

Supposons que $P \nabla \bar{f}_q(e, v_1) > -e$. Alors le vecteur

$$z = \frac{e + P \nabla \bar{f}_q(e, v_1)}{q} (\bar{c}^T e - v_1)$$

est un point dual admissible et le saut de dualité

$$\bar{c}^T e - v_2 = e^T z$$

satisfait

$$\bar{c}^T e - v_2 \leq \frac{n}{q} \left(1 + \frac{\|P \nabla \bar{f}_q(e, v_1)\|}{\sqrt{n}} \right) (\bar{c}^T e - v_1)$$

Démonstration

L'hypothèse

$$P \nabla \bar{f}_q(e, v_1) > -e,$$

nous permet d'assurer le caractère strictement positif de l'expression suivante

$$\frac{e + P \nabla \bar{f}_q(e, v_1)}{q} (\bar{c}^T e - v_1).$$

Autrement dit, $z > 0$.

De plus, le calcul de la projection de z nous donne

$$P z = \frac{e_p + \frac{q}{\bar{c}_p^T e_p - v_1} \bar{c}_p - e_p}{q} (\bar{c}_p^T e_p - v_1) = \bar{c}_p = P \bar{c}.$$

Ainsi, par la proposition 1.3.2. [cfr chap.III], nous pouvons affirmer que z est un point dual admissible.

Le saut de dualité est donné par

$$e^T z = \frac{n + e^T P \nabla \bar{f}_q(e, v_1)}{q} (\bar{c}^T e - v_1).$$

Or, nous avons les relations suivantes

$$\begin{aligned} \left| e^T P \nabla \bar{f}_q(e, v_1) \right| &\leq \|e\| \left\| P \nabla \bar{f}_q(e, v_1) \right\| \\ &\leq \sqrt{n} \left\| P \nabla \bar{f}_q(e, v_1) \right\|. \end{aligned}$$

Il s'en suit que

$$e^T z \leq \frac{n}{q} \left(1 + \frac{\left\| P \nabla \bar{f}_q(e, v_1) \right\|}{\sqrt{n}} \right) (\bar{c}^T e - v_1). \quad \blacksquare$$

Nous pouvons remarquer que dans ce lemme, le critère de proximité est plus faible que celui défini en début de paragraphe. Nous imposons maintenant la condition stricte sur

$\left\| P \nabla \bar{f}_q(e, v) \right\|$ et nous établissons la règle de mise à jour du paramètre v , la borne inférieure de \hat{v} .

Proposition 3.2.

Soit $v_1 < \hat{v}$ une borne inférieure connue.

Supposons qu'après une transformation "scaling",

$$\left\| P \nabla \bar{f}_q (e, v_1) \right\| \leq \nu \text{ où } \nu \in (0, 1) \text{ et } q > n.$$

$$\text{Si } r \geq (1 + \frac{\nu}{\sqrt{n}}) n,$$

$$\text{alors } v_2 = \bar{c}^T e - \frac{r}{q} (\bar{c}^T e - v_1)$$

est une borne inférieure admissible, c'est-à-dire $v_2 < \hat{v}$.

$$\text{De plus, nous avons } \frac{\bar{c}^T e - v_2}{\bar{c}^T e - v_1} = \frac{r}{q}$$

$$\text{et } \nabla \bar{f}_q (e, v_1) = \nabla \bar{f}_r (e, v_2). \quad (48)$$

Démonstration

Comme $\left\| P \nabla \bar{f}_q (e, v_1) \right\| \leq \nu$, chaque composante de $P \nabla \bar{f}_q (e, v_1)$ est comprise entre $-\nu$ et ν où $\nu \in (0, 1)$.

Nous vérifions la condition

$$P \nabla \bar{f}_q (e, v_1) > -e.$$

Nous pouvons ainsi appliquer le lemme 3.1. qui nous assure l'existence d'une borne v_2 ,

$$v_2 = \bar{c}^T e - e^T z \leq \hat{v},$$

avec $z = \frac{e + P \nabla \bar{f}_q (e, v_1)}{q} (\bar{c}^T e - v_1)$ et satisfaisant la relation

$$\frac{\bar{c}^T e - v_2}{\bar{c}^T e - v_1} \leq \frac{n}{q} \left(1 + \frac{\left\| P \nabla \bar{f}_q (e, v_1) \right\|}{\sqrt{n}} \right).$$

Nous pouvons encore majorer ce rapport de telle manière

$$\frac{\bar{c}^T \bar{e} - \bar{v}_2}{\bar{c}^T \bar{e} - \bar{v}_1} \leq \frac{n}{q} \left(1 + \frac{\nu}{\sqrt{n}} \right) \quad (49)$$

puisque

$$\left\| P \nabla \bar{f}_q(e, v_1) \right\| \leq \nu.$$

Or, par hypothèse

$$r \geq \left(1 + \frac{\nu}{\sqrt{n}} \right) n.$$

Dès lors, nous obtenons

$$\frac{\bar{c}^T \bar{e} - \bar{v}_2}{\bar{c}^T \bar{e} - \bar{v}_1} \leq \frac{r}{q} = \frac{\bar{c}^T \bar{e} - \bar{v}_2}{\bar{c}^T \bar{e} - \bar{v}_1}$$

$$\text{car } \bar{v}_2 = \bar{c}^T \bar{e} - \frac{r}{q} (\bar{c}^T \bar{e} - \bar{v}_1).$$

Ainsi, nous avons

$$\bar{c}^T \bar{e} - \bar{v}_2 \leq \bar{c}^T \bar{e} - \bar{v}_2$$

et puisque $\bar{v}_2 \leq \hat{\bar{v}}$, par le lemme 3.1., nous obtenons

$$\bar{v}_2 < \bar{v}_2 \leq \hat{\bar{v}}.$$

Par ailleurs, la définition de \bar{v}_2 nous conduit à

$$\frac{\bar{c}^T \bar{e} - \bar{v}_2}{\bar{c}^T \bar{e} - \bar{v}_1} = \frac{r}{q}$$

et ainsi

$$\frac{q}{\bar{c}^T \bar{e} - \bar{v}_1} = \frac{r}{\bar{c}^T \bar{e} - \bar{v}_2}$$

ou encore

$$\frac{q}{\bar{c}^T \bar{e} - \bar{v}_1} \bar{c} - e = \frac{r}{\bar{c}^T \bar{e} - \bar{v}_2} \bar{c} - e.$$

Autrement dit, nous vérifions la dernière assertion de la proposition :

$$\nabla \bar{f}_q(e, v_1) = \nabla \bar{f}_r(e, v_2). \quad \blacksquare$$

Proposition 3.3.

La meilleure croissance de la borne inférieure est atteinte lorsque

$$r = \left(1 + \frac{\nu}{\sqrt{n}}\right) n \text{ où } \nu \in (0, 1).$$

Démonstration

En vertu de la proposition 3.2., nous savons que

$$r \geq \left(1 + \frac{\nu}{\sqrt{n}}\right) n$$

et

$$\frac{\frac{-T}{c} e^{-v_2}}{\frac{-T}{c} e^{-v_1}} \leq \frac{n}{q} \left(1 + \frac{\nu}{\sqrt{n}}\right) \text{ par (49)}$$

Prenons $r = \left(1 + \frac{\nu}{\sqrt{n}}\right) n$ et $r' > \left(1 + \frac{\nu}{\sqrt{n}}\right) n$.

En utilisant (49), nous obtenons les relations suivantes :

$$\frac{\frac{-T}{c} e^{-v_2}}{\frac{-T}{c} e^{-v_1}} \leq \frac{r}{q} < \frac{r'}{q}.$$

En outre, en considérant les valeurs de r et de r' ,

$$\frac{r}{q} = \frac{\frac{-T}{c} e^{-v_2}}{\frac{-T}{c} e^{-v_1}} \text{ et } \frac{r'}{q} = \frac{\frac{-T}{c} e^{-v_2'}}{\frac{-T}{c} e^{-v_1}}$$

Nous avons ainsi

$$\frac{-T}{c} e^{-v_2} \leq \frac{-T}{c} e^{-v_2} < \frac{-T}{c} e^{-v_2'}$$

ou encore,

$$v_2' < v_2 \leq \hat{v}_2 < \hat{v}.$$

La meilleure croissance de la borne est atteinte en v_2 , c'est-à-dire lorsque

$$r = (1 + \frac{\nu}{\sqrt{n}}) n. \quad \blacksquare$$

4. Un algorithme de complexité $O(nL)$

4.1. L'algorithme

Nous particularisons l'algorithme 5.1 [cfr chapitre IV] à la fonction potentielle. Nous effectuons néanmoins un petit changement de notation : le compteur d'itérations, k , est incrémenté à chaque itération interne et non lorsqu'un point proche de la trajectoire centrale est obtenu.

Algorithme 4.1.1

Soit $q > n + \sqrt{n}$,

$$\varepsilon \in (0, 1]$$

$$v_0 < \hat{v}$$

$$x^0 \in \overset{\circ}{S}$$

$$k = 0$$

Répéter jusqu'à ce que $c^T x^k - v^k < 2^{-L}$

1) "SCALING"

- Appliquer la transformation de coordonnées :

$$X_k^{-1} : \mathcal{R}^n \rightarrow \mathcal{R}^n$$

$$x \rightarrow X_k^{-1} x = y$$

Ainsi x^k est transporté sur le point e .

- Noter $\bar{A} = A X_k$, $\bar{c} = X_k c$

$$\text{et } \bar{f}_q(e, v) = f_q(x^k, v)$$

- Calculer les vecteurs projetés \bar{c}_p et e_p .

2) DIRECTION

$$\bar{h} = -P \nabla \bar{f}_q(e, v_k) = \frac{-q}{c^T e - v_k} \bar{c}_p + e_p$$

3) MISE A JOUR DE v_k

$$\text{Si } \|\bar{h}\| < \varepsilon$$

$$\text{Alors } r = n \left(1 + \frac{\|\bar{h}\|}{\sqrt{n}}\right)$$

$$v_{k+1} = c^T e - \frac{r}{q} (c^T e - v_k)$$

$$\bar{h} = - \frac{q}{c^T e - v_{k+1}} \bar{c}_p + e_p$$

$$\text{Sinon } v_{k+1} = v_k.$$

4) RECHERCHE LINEAIRE

$$\bar{\lambda} = \operatorname{argmin} \{ \bar{f}_q(e + \lambda \bar{h}, v_{k+1}) \mid \lambda > 0, e + \lambda \bar{h} > 0 \}$$

$$y = e + \bar{\lambda} \bar{h}$$

5) " SCALING " inverse :

Retourner dans l'espace d'origine par la transformation
" scaling " inverse,

$$x^{k+1} = X_k y$$

$$6) k = k + 1$$

Remarquons que le choix de $q > n + \sqrt{n}$ garantit une diminution du saut de dualité à chaque mise à jour de la borne inférieure v .

Pour se rendre compte de ce fait, plaçons-nous à l'itération k . La borne v_k est mise à jour. Nous constatons alors que

$$\frac{\frac{-T}{c} e - v_{k+1}}{\frac{-T}{c} e - v_k} = \frac{n}{q} \left(1 + \frac{\|h\|}{\sqrt{n}} \right) \text{ et } \|h\| < \varepsilon.$$

Dès lors, nous obtenons

$$\begin{aligned} \frac{\frac{-T}{c} e - v_{k+1}}{\frac{-T}{c} e - v_k} &< \frac{n}{q} \left(1 + \frac{\varepsilon}{\sqrt{n}} \right) \\ &= \frac{n + \sqrt{n} \varepsilon}{q} \end{aligned}$$

Or $q > n + \sqrt{n}$, il suit que

$$\begin{aligned} \frac{n + \sqrt{n} \varepsilon}{q} &< \frac{n + \sqrt{n} \varepsilon}{n + \sqrt{n}} \\ &\leq \frac{n + \sqrt{n}}{n + \sqrt{n}} \quad (\varepsilon \leq 1) \\ &= 1 \end{aligned}$$

Nous concluons que

$$\frac{-T}{c} e - v_{k+1} < \frac{-T}{c} e - v_k.$$

C'est-à-dire, le saut de dualité diminue après la mise à jour de v_k .

4.2. Etude de la complexité

Nous montrons tout d'abord qu'à chaque itération, la valeur de la fonction potentielle diminue d'au moins une valeur constante qui ne dépend pas de l'itération.

Nous avons deux possibilités à envisager :

◆ A. $\left\| P \nabla \bar{f}_q (e, v) \right\| = \left\| \bar{h} \right\| \geq \varepsilon$

Dans ce cas, la fonction potentielle diminue le long de la direction \bar{h} d'au moins une constante qui dépend seulement de ε .

Nous prouvons ceci pour une valeur spécifique de ε : $\varepsilon = 1$, [1].

Lemme 4.2.1.

Soient $\lambda > 0$ et une direction $h \in \mathbb{R}^n$ tels que $\lambda \|h\|_\infty \leq 0,5$

Alors,

$$f_q (e + \lambda h, v) \leq f_q (e, v) + \lambda \nabla f_q (e, v)^T h + 2 \lambda^2 \|h\|^2.$$

Démonstration

Considérons la fonction f_2 définie comme suit :

$$y \in \mathbb{R}^n, y > 0 \rightarrow f_2 (y) = \sum_{i=1}^n \log y_i .$$

Considérons aussi un point $y \in \mathbb{R}^n, y > 0$, une direction $h \in \mathbb{R}^n$ et un pas $\lambda > 0$ tels que

$$y + \lambda h > 0.$$

Nous avons alors par définition de f_2 :

$$f_2 (y + \lambda h) = \sum_{i=1}^n \log (y_i + \lambda h_i).$$

Développons $\log (y_i + \lambda h_i)$ autour de chaque y_i , pour obtenir:

$$f_2 (y + \lambda h) = \sum_{i=1}^n (\log y_i + \lambda h_i^T y_i^{-1} + o (y_i, \lambda h_i))$$

$$\text{où } o (y_i, \lambda h_i) = - \frac{\lambda^2 h_i^2}{2 (y_i + \Theta_i)^2} \quad (\Theta_i \in (0, \lambda h_i))$$

et donc $f_2(y + \lambda h) = \sum_{i=1}^n \log y_i + \lambda h^T y^{-1} + \sum_{i=1}^n o(y_i, \lambda h_i)$.

Il suit que

$$f_2(y + \lambda h) \geq \sum_{i=1}^n \log y_i + \lambda h^T y^{-1} - \sum_{i=1}^n \frac{\lambda^2 h_i^2}{2(y_i - |\lambda h_i|)^2}.$$

En particulier, pour $y_i - \lambda |h_i| \geq 0,5$ ($i = 1, \dots, n$), nous obtenons

$$f_2(y + \lambda h) \geq f_2(y) + \lambda h^T y^{-1} - 2\lambda^2 \|h\|^2.$$

Considérons finalement la fonction potentielle $f_q(y, v)$:

Le premier terme, $f_1(y) = q \log(c^T y - v)$, est une fonction strictement concave et donc est bornée supérieurement par l'approximation linéaire.

En d'autres termes,

$$f_1(y + \lambda h) \leq f_1(y) + \lambda \nabla f_1(y)^T h.$$

Rassemblons les relations obtenues pour f_1 et f_2 et nous obtenons ainsi:

Si $y_i - \lambda |h_i| \geq 0,5$ ($i = 1, \dots, n$),

$$\begin{aligned} f_q(y + \lambda h) &= f_1(y + \lambda h) - f_2(y + \lambda h) \\ &\leq f_q(y, v) + \lambda \nabla f_q(y)^T h + 2\lambda^2 \|h\|^2. \end{aligned}$$

En posant $y = e$, cette propriété devient

Si $\lambda \|h\|_\infty \leq 0,5$,

$$f_q(e + \lambda h) = f_q(e, v) + \lambda \nabla f_q(e)^T h + 2\lambda^2 \|h\|^2 \quad \blacksquare$$

Note :

Cette propriété est valable pour toute fonction de même type que la fonction potentielle et donc en particulier pour \bar{f}_q qui ne diffère de f_q que par une constante.

Proposition 4.2.2.

$$\text{Si } \|\bar{h}\| \geq 1,$$

$$\text{Alors } \bar{f}_q \left(e + \frac{0,3}{\|\bar{h}\|} \bar{h}, v \right) - \bar{f}_q (e, v) < -0,1$$

Démonstration

Puisque

$$\frac{0,3}{\|\bar{h}\|} \|\bar{h}\|_\infty \leq \frac{0,3}{\|\bar{h}\|} \|\bar{h}\| \leq 0,5,$$

nous pouvons appliquer le lemme 4.2.1., avec $\lambda = \frac{0,3}{\|\bar{h}\|}$ et $h = \bar{h}$, qui établit la relation suivante :

$$\bar{f}_q \left(e + \frac{0,3}{\|\bar{h}\|} \bar{h}, v \right) \leq \bar{f}_q (e, v) + \frac{0,3}{\|\bar{h}\|} \nabla \bar{f}_q(e, v)^T \bar{h} + 2 \left(\frac{0,3}{\|\bar{h}\|} \right)^2 \|\bar{h}\|^2$$

Comme $\bar{h} \in N(\bar{A})$, nous avons

$$\begin{aligned} \nabla \bar{f}_q (e, v)^T \bar{h} &= P \nabla \bar{f}_q (e, v)^T \bar{h} \\ &= -\|\bar{h}\|^2 \end{aligned}$$

ce qui nous permet d'écrire l'inégalité suivante :

$$\begin{aligned} \bar{f}_q \left(e + \frac{0,3}{\|\bar{h}\|} \bar{h}, v \right) - \bar{f}_q (e, v) &\leq \frac{-0,3}{\|\bar{h}\|} \|\bar{h}\|^2 + 2 (0,3)^2 \\ &= -0,3 \|\bar{h}\| + 0,18. \end{aligned}$$

Or $\|\bar{h}\| \geq 1$ par hypothèse. Nous obtenons finalement

$$\begin{aligned} \bar{f}_q(e + \frac{0,3}{\|\bar{h}\|} \bar{h}, v) - \bar{f}_q(e, v) &\leq -0,3 + 0,18 \\ &< -0,1 \end{aligned}$$

c'est-à-dire la thèse. ■

♦ B. $\|P \nabla \bar{f}_q(e, v)\| = \|\bar{h}\| < \varepsilon.$

Supposons que $\|P \nabla \bar{f}_q(e, v)\| = \|\bar{h}\| < \varepsilon.$

Dans ce cas, la borne inférieure est mise à jour de telle manière que

$$\frac{\bar{c}^T e - v_{k+1}}{\bar{c}^T e - v_k} < \frac{n}{q} \left(1 + \frac{\|\bar{h}\|}{\sqrt{n}}\right)$$

Cependant $\|\bar{h}\| < \varepsilon \leq 1$ et en prenant le logarithme de chaque membre de l'inégalité, il vient que

$$\log(\bar{c}^T e - v_{k+1}) - \log(\bar{c}^T e - v_k) < \log\left(\frac{n}{q} \left(1 + \frac{1}{\sqrt{n}}\right)\right).$$

En outre, supposons que

$$\frac{q}{n} \geq e^{\delta/q} \left(1 + \frac{1}{\sqrt{n}}\right).$$

Autrement dit,

$$\frac{n}{q} \leq e^{-\delta/q} \left(1 + \frac{1}{\sqrt{n}}\right)^{-1}. \quad (50)$$

Nous avons alors en prenant à nouveau le logarithme de chaque membre de (50),

$$\log \frac{n}{q} \left(1 + \frac{1}{\sqrt{n}}\right) \leq \log e^{-\delta/q} = -\frac{\delta}{q}.$$

Par conséquent, nous obtenons

$$q \log (c^T e - v_{k+1}) - q \log (c^T e - v_k) < -\delta$$

et donc

$$\bar{f}_q(e, v_{k+1}) - \bar{f}_q(e, v_k) < -\delta.$$

Remarque

La condition (50) est obtenue facilement en prenant par exemple $q = 2n$ et $\delta = 0,5$ pour $n \geq 2$.

Nous concluons qu'à chaque itération, la fonction potentielle diminue d'au moins une constante valant le minimum entre 0,1 et δ .

Nous sommes maintenant en mesure de montrer que l'algorithme a une complexité de l'ordre de $O(nL)$ itérations, comme il était annoncé.

Lemme 4.2.3.

Si $\overset{\circ}{S} = \{x \in \mathbb{R}^n / Ax = b \text{ et } x > 0\}$ est borné,
la fonction de pénalité f_2 définie par

$$x \in \overset{\circ}{S} \rightarrow \sum_{i=1}^n \log x_i$$

admet une valeur maximale, M .

Démonstration

La thèse revient à prouver l'existence d'une solution au problème

$$\left[\begin{array}{l} \text{Maximiser } \sum_{i=1}^n \log x_i \\ \text{sous } \begin{array}{l} Ax = b \\ x > 0 \end{array} \end{array} \right.$$

Considérons $T = \{x \in \overset{\circ}{S} \mid \sum_{i=1}^n \log x_i \geq \sum_{i=1}^n \log \bar{x}_i\}$

où $\bar{x} \in \overset{\circ}{S}$.

Nous pouvons montrer que

i) T est borné

T est inclus dans $\overset{\circ}{S}$ et $\overset{\circ}{S}$ est borné, donc T est borné.

ii) T est fermé

Prenons une suite de points dans T, (x^k) , convergente vers un élément x.

Montrons que ce dernier appartient à T.

Autrement dit, vérifions que

$$\sum_{i=1}^n \log x_i \geq \sum_{i=1}^n \log \bar{x}_i$$

La suite x^k appartient à T, donc $\sum_{i=1}^n \log x_i^k \geq \sum_{i=1}^n \log \bar{x}_i \quad \forall k$

Par passage à la limite, lorsque k tend vers l'infini, nous obtenons

$$\sum_{i=1}^n \log x_i^k \rightarrow \sum_{i=1}^n \log x_i \quad \text{et} \quad \sum_{i=1}^n \log x_i \geq \sum_{i=1}^n \log \bar{x}_i.$$

De plus, chaque élément de la suite (x^k) appartient à T, d'où $A x^k = b \quad \forall k$.

Par passage à la limite, quand k tend vers l'infini, nous avons $A x = b$ et $x > 0$.

Nous avons ainsi vérifié les conditions d'appartenance à T pour x.

Nous pouvons en déduire que la fonction continue f_2 atteint sa borne supérieure sur T, soit x^* cette borne.

Nous pouvons montrer que x^* est solution du problème :

En effet, si $x \notin T$ et $x \in \overset{\circ}{S}$, nous avons la relation suivante

$$\sum_{i=1}^n \log x_i < \sum_{i=1}^n \log \bar{x}_i \leq \sum_{i=1}^n \log x_i^*$$

et donc x n'est pas une borne supérieure.

La thèse est ainsi montrée en notant $\sum_{i=1}^n \log x_i^* = M$ ■

Lemme 4.2.4.

Si $f_q(x^k, v_k) \leq f_q(x^{k-1}, v_{k-1}) - \alpha$
où $k = 1, 2, \dots$

Alors le problème est résolu en au plus $O(\frac{q}{\alpha} L)$ itérations
(L est la longueur totale des données du problème).

Démonstration

Soit x^0 un point initial.

Par l'hypothèse, nous déduisons que

$$f_q(x^k, v_k) \leq f_q(x^0, v_0) - k \alpha$$

ou de façon équivalente,

$$q \log(c^T x^k - v_k) - \sum_{i=1}^n \log x_i^k \leq f_q(x^0, v_0) - k \alpha.$$

Il suit du lemme 4.2.3 que

$$q \log(c^T x^k - v_k) \leq M + f_q(x^0, v_0) - k \alpha$$

ou encore

$$\log(c^T x^k - v_k) \leq \frac{M + f_q(x^0, v_0)}{q} - k \frac{\alpha}{q}. \quad (51)$$

Par ailleurs, nous considérons le problème résolu quand un point x^k vérifie la condition d'arrêt de l'algorithme, c'est-à-dire,

$$c^T x^k - v_k < 2^{-L}$$

ou en prenant le logarithme de chaque membre

$$\log(c^T x^k - v_k) < -L$$

$$(\log(c^T x^k - v_k) < \log 2^{-L} = -L \log 2 < -L).$$

Par (51), cette condition est garantie pour tout k tel que

$$\frac{M + f_q(x^o, v_o)}{q} - k \frac{\alpha}{q} \leq -L.$$

Dès lors, le nombre d'itérations pour résoudre le problème doit satisfaire la relation

$$k \geq \frac{q}{\alpha} \left(L + \frac{M + f_q(x^o, v_o)}{q} \right).$$

Or $\frac{M + f_q(x_o, v_o)}{q} \leq O(L)$ et donc,

$$k = O\left(\left(\frac{q}{\alpha}\right) L\right). \quad \blacksquare$$

Proposition 4.2.5.

*Si $q = O(n)$,
l'algorithme a une borne de complexité de $O(nL)$ itérations.*

Démonstration

Nous savons maintenant qu'à chaque itération, la valeur de la fonction potentielle diminue d'une constante $K = \min(0,1, \delta)$.

En vertu du lemme 4.2.4., l'algorithme se termine en $O\left(\left(\frac{q}{K}\right) L\right)$ itérations.

Comme $q = O(n)$, nous obtenons une complexité de $O(nL)$ comme annoncé. \blacksquare

5. Un algorithme de complexité $O(\sqrt{n} L)$

5.1. L'algorithme

Nous répétons l'algorithme 4.1.1 en prenant q et r tels que

$$r = 2n, q > r \text{ et } \varepsilon = 0,03.$$

Algorithme 5.1.1.

Soit $q > r = 2n$,

$$v_o < \hat{v},$$

$$x^o \in \overset{\circ}{S}$$

$$k = 0$$

Répéter jusqu'à ce que $c^T x^k - v_k < 2^{-L}$

1) " SCALING " : $X_k^{-1} : \mathcal{R}^n \rightarrow \mathcal{R}^n$

$$x \rightarrow X_k^{-1} x = y$$

2) DIRECTION :

$$\bar{h} = -P \nabla \bar{f}_q(e, v_k).$$

3) MISE A JOUR DE v_k

$$\text{Si } \|\bar{h}\| < 0,03$$

$$\text{Alors } v_{k+1} = \bar{c}^T e - \frac{r}{q} (\bar{c}^T e - v_k)$$

$$\bar{h} = \frac{-q}{\bar{c}^T e - v_{k+1}} \bar{c}_p^T + e_p$$

$$\text{sinon } v_{k+1} = v_k$$

4) RECHERCHE LINEAIRE :

$$\bar{\lambda} = \operatorname{argmin} \{ \bar{f}_q(e + \lambda \bar{h}, v_{k+1}) / \lambda > 0, e + \lambda \bar{h} > 0 \}$$

$$y = e + \bar{\lambda} \bar{h}$$

5) " SCALING " inverse :

$$x^{k+1} = X_k y$$

6) $k = k + 1$.

5.2. Etude de la complexité

Nous abordons l'étude de la complexité de l'algorithme en montrant à nouveau qu'à chaque itération la valeur de la fonction potentielle diminue.

A cet effet, il est nécessaire de rappeler quelques propriétés relatives à la fonction logarithmique et à la fonction barrière logarithmique.

Lemme 5.2.1.

$\forall \lambda \in (-1, 1)$, nous avons

$$1) \log(1 + \lambda) \leq \lambda$$

$$2) \log(1 + \lambda) \geq \lambda - \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}$$

$$3) \log(1 + \lambda) \leq \lambda - \frac{\lambda^2}{2} + \frac{\lambda^3}{3}$$

Démonstration

Les propriétés 1) et 2) ont été montrées dans le chapitre III [lemme 2.2.1.] et la troisième découle immédiatement du développement en séries de Taylor de la fonction logarithmique autour de 1.

Lemme 5.2.2.

Soient $\lambda \in (-1, +1)$ et h un vecteur de \mathbb{R}^n tel que $\|h\| = 1$.

Nous pouvons montrer

$$1) p(e + \lambda h) \geq -\lambda e^T h = \lambda \nabla p(e)^T h$$

$$2) p(e + \lambda h) \leq \lambda \nabla p(e)^T h + \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}$$

$$3) p(e + \lambda h) \geq \lambda \nabla p(e)^T h + \frac{\lambda^2}{2} - \frac{\lambda^3}{3}$$

Démonstration

Les propriétés 1) et 2) ont été montrées dans le chapitre III (Proposition 2.2.2.).
Montrons la troisième propriété :

Comme $\|\lambda h\|_\infty \leq |\lambda| \|h\| < 1$,

chacune des composantes λh_i est comprise entre -1 et +1.

Nous pouvons ainsi utiliser la propriété 3 du lemme 5.2.1., pour obtenir

$$p(e + \lambda h) \geq -\lambda \sum_{i=1}^n h_i + \frac{\lambda^2}{2} \sum_{i=1}^n h_i^2 - \frac{\lambda^3}{3} \sum_{i=1}^n h_i^3.$$

$$\text{Or } \sum_{i=1}^n h_i^3 \leq \sum_{i=1}^n |h_i|^3 \leq \max |h_i| \sum_{i=1}^n h_i^2 = \|h\|_\infty \sum_{i=1}^n h_i^2 \leq 1.$$

Nous en déduisons que

$$p(e + \lambda h) \geq -\lambda e^T h + \frac{\lambda^2}{2} - \frac{\lambda^3}{3}. \quad \blacksquare$$

Proposition 5.2.3.

Soit $h \in \text{Ker}(A)$, une direction telle que $\|h\| = 1$.

Supposons que $q \geq 2n$.

Alors nous pouvons montrer les deux assertions suivantes :

$$1) \text{ Si } \nabla f_q(e, v)^T h \leq -0,03$$

$$\text{alors } f_q(e + 0,04 h, v) \leq f_q(e, v) - 0,00024$$

$$2) \text{ Si } \|P \nabla f_q(e, v)\| \leq 0,03$$

$$\text{alors } f_q(e, v) - f_q(x_q(v), v) \leq 0,0011$$

Démonstration

Considérons le premier terme de la fonction potentielle,

$$f_1(x) = q \log(c^T x - v).$$

Évaluée en $e + \lambda h$, elle devient

$$f_1(e + \lambda h) = q \log(c^T e + c^T \lambda h - v).$$

Faisons la transformation suivante

$$(c^T e + c^T \lambda h - v) = (c^T e - v) \left(1 + \lambda \frac{c^T h}{c^T e - v}\right).$$

pour obtenir

$$f_1(e + \lambda h) = q \log(c^T e - v) + q \log\left(1 + \lambda \frac{c^T h}{c^T e - v}\right).$$

Définissons maintenant

$$a = \frac{c^T h}{c^T e - v}.$$

Notons que $a = \frac{1}{q} \nabla f_1(e)^T h$. Nous obtenons ainsi

$$f_1(e + \lambda h) = f_1(e) + q \log(1 + \lambda a)$$

Après ces considérations, montrons la première assertion de la proposition :

1) En utilisant la première propriété du lemme 5.2.1., nous constatons que

$$f_1(e + \lambda h) \leq f_1(e) + q \lambda a$$

et par définition de a ,

$$f_1(e + \lambda h) = f_1(e) + \lambda \nabla f_1(e)^T h.$$

D'autre part,

$$f_q(e + \lambda h, v) = f_1(e + \lambda h) + p(e + \lambda h).$$

Dès lors,

$$f_q(e + \lambda h, v) \leq f_1(e) + \lambda \nabla f_1(e)^T h + p(e + \lambda h).$$

Appliquons la propriété 2) du lemme 5.2.2. de sorte que

$$\begin{aligned} f_q(e + \lambda h, v) &\leq f_1(e) + \lambda \nabla f_1(e)^T h + \lambda \nabla p(e)^T h + \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|} \\ &= f_q(e, v) + \lambda \nabla f_q(e, v)^T h + \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}. \end{aligned}$$

Par hypothèse, $\nabla f_q(e, v)^T h \leq -0,03$. Il suit que

$$f_q(e + \lambda h, v) \leq f_q(e, v) - 0,03 \lambda + \frac{\lambda^2}{2} \frac{1}{1 - |\lambda|}.$$

En substituant, $\lambda = 0,04$, dans cette dernière inégalité, nous obtenons la thèse.

Il nous reste à montrer la deuxième assertion :

2) Définissons la direction normalisée au point central,

$$h = \frac{x_q(v) - e}{\|x_q(v) - e\|}$$

Notons $x_q(v) = e + \lambda h$, $\lambda > 0$.

La preuve de l'assertion se divise en quatre étapes :

i) Evaluons pour un temps la valeur de a :

Nous avons défini

$$a = \frac{1}{q} \nabla f_1(e)^T h.$$

Ainsi, nous exprimons

$$\begin{aligned} \nabla f_q(e, v)^T h &= \nabla f_1(e)^T h + \nabla p(e)^T h \\ &= a q - e^T h. \end{aligned}$$

Or, nous savons que

$$\begin{aligned} |\nabla f_q(e, v)^T h| &= |P \nabla f_q(e, v)^T h| \\ &\leq \|P \nabla f_q(e, v)\| \|h\| \\ &\leq 0,03. \end{aligned} \quad (52)$$

Il suit que

$$|a q - e^T h| \leq 0,03 \text{ et } |a| q \leq |e^T h| + 0,03.$$

Examinons $|e^T h|$:

Par le lemme 3.2.1. [chap. III], nous pouvons affirmer que la direction normalisée a au moins une composante positive et négative.
(S est compact).

Prenons s.p.d.g. $h_n < 0$ et $h_{n-1} > 0$.

$$\text{Ainsi } |e^T h| \leq \|(h_1, \dots, h_{n-2}, h_{n-1} + h_n)\| \sqrt{n-1}.$$

Développons cette borne en considérant les relations suivantes

$$\begin{aligned}
 \|(h_1, \dots, h_{n-2}, h_{n-1} + h_n)\|^2 &= h_1^2 + \dots + h_{n-2}^2 + (h_{n-1} + h_n)^2 \\
 &= 1 - (h_{n-1}^2 + h_n^2) + (h_{n-1} + h_n)^2 \\
 &< 1 + 2 h_{n-1} h_n \\
 &< 1 \quad (h_{n-1} \text{ et } h_n \text{ sont de signes opposés}).
 \end{aligned}$$

nous avons donc

$$|e^T h| \leq \sqrt{n-1}.$$

Par conséquent,

$$|a| \leq \frac{\sqrt{n-1} + 0,03}{q}.$$

De plus, nous pouvons observer que le membre de droite de l'inégalité décroît avec n , pour $q = 2n$.

Comme $n \geq 2$, nous avons pour $q \geq 2n$,

$$|a| \leq \frac{\sqrt{n-1} + 0,03}{2n} \leq 0,26$$

et aussi

$$q a^2 \leq \frac{(\sqrt{n} + 0,03)^2}{2n} \leq 0,52.$$

ii) Nous avons pu montrer à l'étape (i) que $|a| < 1$.

Nous sommes ainsi en mesure d'appliquer la deuxième propriété du lemme 5.2.1., pour obtenir les relations suivantes :

$$f_1(e + \lambda h) = f_1(e) + q \log(1 + \lambda a)$$

$$\geq f_1(e) + q \left(\lambda a - \frac{\lambda^2 a^2}{2} \frac{1}{1 - |\lambda a|} \right)$$

$$= f_1(e) + \lambda \nabla f_1(e)^T h - q \frac{\lambda^2 a^2}{2} \frac{1}{1 - |\lambda a|}$$

En appliquant ensuite la propriété 3 du lemme 5.2.2., nous établissons

$$f_q(e + \lambda h, v) = f_1(e + \lambda h) + p(e + \lambda h)$$

$$\geq f_1(e) + \lambda \nabla f_1(e)^T h - q \frac{\lambda^2 a^2}{2} \frac{1}{1 - |\lambda a|} + \lambda \nabla p(e)^T h + \frac{\lambda^2}{2} - \frac{\lambda^3}{3}$$

ou encore

$$f_q(e + \lambda h, v) \geq f_q(e, v) + \lambda \nabla f_q(e, v)^T h + \frac{\lambda^2}{2} - \frac{\lambda^3}{3} - q \frac{\lambda^2 a^2}{2} \frac{1}{1 - |\lambda a|}.$$

Nous pouvons simplifier cette inégalité en considérant

$$\nabla f_q(e, v)^T h \geq -0,03 \quad \text{par (52)}$$

$$\text{et } q a^2 \leq 0,52 \quad \text{par (i).}$$

Ainsi, elle devient

$$f_q(e + \lambda h, v) - f_q(e, v) \geq -0,03 \lambda + \frac{\lambda^2}{2} - \frac{\lambda^3}{3} - 0,26 \frac{\lambda^2}{1 - 0,26 \lambda}.$$

$$\text{Notons } \rho(\lambda) \equiv -0,03 \lambda + \frac{\lambda^2}{2} - \frac{\lambda^3}{3} - 0,26 \frac{\lambda^2}{1 - 0,26 \lambda}.$$

Dans la troisième étape, nous étudions $\rho(\lambda)$.

iii) Nous pouvons montrer les propriétés suivantes concernant $\rho(\lambda)$:

- $\rho(\lambda)$ a un minimum local en $\lambda = 0,073$.

En effet,

$$\rho'(\lambda) = -0,03 + \lambda - \lambda^2 - 0,26 \left(\frac{2\lambda - 0,26\lambda^2}{(1 - 0,26\lambda)^2} \right) \quad \text{et} \quad \rho'(0,073) = 0.$$

De plus $\rho''(\lambda) > 0$ pour $\lambda = 0,073$.

Dès lors $\lambda = 0,073$ est un minimum local de $\rho(\lambda)$.

- En évaluant $\rho(\lambda)$ en $\lambda = 0,18$, nous constatons que $\rho(\lambda) > 0$.
- $\forall \lambda \in [0, 0,18], \rho(\lambda) \geq -0,0012$
 $\rho(0) = 0$ et $\rho(0,18) > 0$.
 De plus $0,073$, un minimum local de $\rho(\lambda)$, appartient à l'intervalle $[0, 0,18]$ et $\rho(0,073) \cong -0,0011$.
 Nous en déduisons que $\rho(\lambda) \geq -0,0012$ pour $\lambda \in [0, 0,18]$.

Nous pouvons à présent conclure :

iv) Nous avons établi que

$$f_q(e + \lambda h, v) - f_q(e, v) \geq \rho(\lambda).$$

Comme $\rho(\lambda)$ est une fonction positive en $\lambda = 0,018$, le membre de gauche de l'inégalité est strictement positif en $\lambda = 0,018$.

De plus $f_q(e + \lambda h, v) - f_q(e, v)$ est unimodale (Propriété 3 de f_q [cfr chap.III]) et décroît en $\lambda = 0$.

Il en découle qu'il existe un minimum absolu $\bar{\lambda} \in [0, 0,18]$ appartenant à l'intervalle et vérifiant

$$f_q(e + \bar{\lambda} h, v) - f_q(e, v) \geq - 0,0012$$

ou encore

$$f_q(x_q(v), v) - f_q(e, v) \geq - 0,0012. \quad \blacksquare$$

Nous pouvons à présent prouver le résultat principal montrant l'amélioration dans le saut de dualité, entre deux mises à jour de la borne inférieure. La preuve de la complexité suivra alors de manière triviale.

Dans l'analyse de complexité, nous utiliserons $q = 2n + \gamma / \sqrt{n}$ où $\gamma \geq 1$.

Lemme 5.2.4.

Soit $q = 2n + \gamma \sqrt{n}$ où $\gamma \geq 1$ et $\delta = 0,00024$. Considérons k et $k + J$, deux itérations consécutives au cours desquelles la borne inférieure est mise à jour.

Nous montrons la relation suivante :

$$\log(c^T x^{k+j} - v_{k+j}) \leq \log(c^T x^k - v_k) - j \frac{\delta}{\gamma \sqrt{n}}$$

où $1 \leq j \leq J$.

Démonstration

Considérons l'itération k . Par hypothèse, la borne inférieure v_k doit être mise à jour.

Après une transformation " scaling " autour de x^k , nous devons avoir

$$\left\| P \nabla \bar{f}_q(e, v_k) \right\| \leq 0,03.$$

De plus, par la proposition 3.2., (48), nous avons par construction

$$P \nabla \bar{f}_r(e, v_{k+j}) = P \nabla \bar{f}_q(e, v_k). \\ \text{où } 1 \leq j \leq J$$

Dès lors, nous avons aussi

$$\left\| P \nabla \bar{f}_r(e, v_{k+j}) \right\| \leq 0,03.$$

En appliquant la proposition 5.2.3., nous pouvons montrer deux inégalités :

$$1) r \log (c^T x^{k+j} - v_{k+j}) + p(x^{k+j}) \geq r \log (c^T x^k - v_{k+j}) + p(x^k) - \Delta \quad (53) \\ \text{où } \Delta = 0,0011.$$

En effet, comme

$$\left\| P \nabla \bar{f}_r(e, v_{k+j}) \right\| \leq 0,003,$$

la deuxième assertion de la Proposition 5.2.3. implique que

$$\bar{f}_r(e, v_{k+j}) - \bar{f}_r(\bar{x}_r(v_{k+j}), v_{k+j}) \leq 0,0011$$

ou encore

$$f_r(x^k, v_{k+j}) - 0,0011 \leq f_r(x_r(v_{k+j}), v_{k+j}).$$

Puisque $x_r(v_{k+j})$ est le point central associé à v_{k+j} , nous avons

$$f_r(x_r(v_{k+j}), v_{k+j}) \leq f_r(x^{k+j}, v_{k+j}).$$

Dès lors,

$$f_r(x^k, v_{k+j}) - 0,0011 \leq f_r(x^{k+j}, v_{k+j}),$$

c'est-à-dire (53).

$$2) f_q(x^{k+j}, v_{k+j}) \leq f_q(x^k, v_{k+j}) - (j-1)\delta \quad \text{où } 1 \leq j \leq J \quad (54)$$

Procédons par étape :

- $j = 1$:

x^{k+1} est le point obtenu, à l'itération k , après la mise à jour de v_k et une recherche linéaire le long de

$$\bar{h} = - \frac{q}{\bar{c}^T e - v_{k+1}} \bar{c}_p + e_p,$$

$$\text{d'où} \quad f_q(x^{k+1}, v_{k+1}) \leq f_q(x^k, v_{k+1}).$$

- $j = 2$:

$$\text{Posons } \bar{h} = - \frac{P \nabla \bar{f}_q(e, v_{k+1})}{\|P \nabla \bar{f}_q(e, v_{k+1})\|}. \quad \text{Comme } \|\bar{h}\| = \|P \nabla \bar{f}_q(e, v_{k+1})\| \geq 0,03,$$

nous avons

$$\left| P \nabla \bar{f}_q(e, v_{k+1})^T \bar{h} \right| = \|\bar{h}\| \geq 0,03,$$

ou encore

$$- P \nabla \bar{f}_q(e, v_{k+1})^T \bar{h} \geq 0,03.$$

L'assertion 1 de la Proposition 5.2.3. nous assure alors que

$$\bar{f}_q(e + 0,04 \bar{h}, v_{k+1}) \leq \bar{f}_q(e, v_{k+1}) - \delta.$$

$$\text{Comme } \bar{h} = \frac{\bar{h}}{\|\bar{h}\|}, \text{ nous obtenons}$$

$$\begin{aligned} \bar{f}_q(e + \bar{\lambda} \bar{h}, v_{k+1}) &\leq \bar{f}_q\left(e + \frac{0,04}{\|\bar{h}\|} \bar{h}, v_{k+1}\right) \\ &\leq \bar{f}_q(e, v_{k+1}) - \delta \end{aligned}$$

$$\text{où } \bar{\lambda} = \operatorname{argmin} \{ \bar{f}_q(e + \lambda \bar{h}, v_{k+1}) \mid \lambda > 0, e + \lambda \bar{h} > 0 \}.$$

Effectuons une transformation " scaling " inverse, de matrice X_{k+1} , et rappelons-nous que $v_{k+1} = v_{k+2}$. La dernière inégalité peut s'écrire sous la forme suivante,

$$f_q(x^{k+2}, v_{k+2}) \leq f_q(x^{k+1}, v_{k+2}) - \delta.$$

Or $f_q(x^{k+1}, v_{k+1}) \leq f_q(x^k, v_{k+1})$. Par conséquent,

$$f_q(x^{k+2}, v_{k+2}) \leq f_q(x^k, v_{k+2}) - \delta.$$

Par le même raisonnement, nous obtenons successivement

$$f_q(x^{k+3}, v_{k+3}) \leq f_q(x^k, v_{k+3}) - 2\delta,$$

....

et enfin

$$f_q(x^{k+J}, v_{k+J}) \leq f_q(x^k, v_{k+J}) - (J - 1)\delta.$$

C'est-à-dire (54).

En soustrayant les inégalités, soit (54) - (53), et en considérant $r = 2n$ et

$q = 2n + \gamma\sqrt{n}$, nous obtenons

$$\gamma\sqrt{n} \log(c^T x^{k+j} - v_{k+j}) \leq \gamma\sqrt{n} \log(c^T x^k - v_{k+j}) + \Delta - (j - 1)\delta. \quad (55)$$

En outre, à l'itération k , la borne inférieure v_k est mise à jour et le saut de dualité vérifie la relation

$$\frac{c^T x^k - v_{k+j}}{c^T x^k - v_k} = \frac{r}{q} = \frac{2n}{2n + \gamma\sqrt{n}}$$

ou encore, en appliquant la fonction logarithme,

$$\log(c^T x^k - v_{k+j}) = \log(c^T x^k - v_k) + \log \frac{2n}{2n + \gamma\sqrt{n}}.$$

$$\text{Or } \log \frac{2n}{2n + \gamma\sqrt{n}} = -\log \left(1 + \frac{\gamma}{2\sqrt{n}}\right).$$

Nous obtenons ainsi

$$\log(c^T x^k - v_{k+j}) = \log(c^T x^k - v_k) - \log \left(1 + \frac{\gamma}{2\sqrt{n}}\right).$$

Substituons cette dernière expression dans l'inégalité (55). Il vient, après quelques opérations élémentaires, que

$$\log(c^T x^k - v_{k+j}) - \log(c^T x^k - v_k) \leq \frac{\Delta - (j - 1)\delta}{\gamma\sqrt{n}} - \log \left(1 + \frac{\gamma}{2\sqrt{n}}\right).$$

Majorons le membre de droite de l'inégalité en considérant

$$\text{i) } \log \left(1 + \frac{\gamma}{2\sqrt{n}} \right) \geq \frac{0,4}{\sqrt{n}} \quad \text{car } \gamma \geq 1 \text{ et } n \geq 2 ,$$

$$\text{ii) } \Delta - 0,4 \gamma \leq \Delta - 0,4 \leq -\delta \quad \text{car } \gamma \geq 1 .$$

Par conséquent, nous obtenons finalement

$$\begin{aligned} \log (c^T x^k - v_{k+j}) - \log (c^T x^k - v_k) &\leq \frac{1}{\gamma \sqrt{n}} (\Delta - 0,4 \gamma - (j-1) \delta) \quad \text{par (i)} \\ &\leq -j \frac{\delta}{\gamma \sqrt{n}} \quad \text{par (ii).} \end{aligned}$$

Le lemme est ainsi montré. ■

Nous en venons à présent à la preuve de convergence de l'algorithme 5.1.1.
Nous supposons pour cela que $x^\circ = e$ est un point proche de la trajectoire centrale,
Autrement dit,

$$\left\| P \nabla \bar{f}_q(e, v_o) \right\| \leq 0,03.$$

Proposition 5.2.5.

Supposons que l'algorithme 5.1.1. commence avec $x^\circ = e$ et $\left\| P \nabla \bar{f}_q(e, v_o) \right\| \leq 0,03$.

Supposons également que

$$c^T e - v_o < 2^L, r = 2n \text{ et } q = r + \gamma \sqrt{n} \quad (\gamma \geq 1).$$

Nous pouvons alors montrer que l'algorithme 5.1.1. stoppe en au plus $O(\gamma \sqrt{n} L)$ itérations.

Démonstration

Sous les hypothèses de la proposition, la borne inférieure est mise à jour à la première itération.

Appliquons le lemme 5.2.4. de manière récursive à des suites d'itérations consécutives pour lesquelles la borne inférieure est fixée :

Nous obtenons pour tout itération k

$$\log (c^T x^k - v_k) \leq \log (c^T x^o - v_o) - k \frac{\delta}{\gamma \sqrt{n}}.$$

En prenant $k \geq \frac{2 \gamma \sqrt{n} L}{\delta}$, c'est-à-dire $k = O(\gamma \sqrt{n} L)$,
nous vérifions la condition d'arrêt de l'algorithme :

$$\log (c^T x^k - v_k) \leq \log (c^T x^o - v_o) - 2 L.$$

Or $c^T x_o - v_o < 2^L$, par hypothèse, nous en déduisons que

$$\log (c^T x^k - v_k) < -L$$

ou encore

$$c^T x^k - v_k < 2^{-L}.$$

L'algorithme 5.1.1. se termine ainsi après un nombre fini d'itérations de l'ordre de $O(\gamma \sqrt{n} L)$.

Si $\gamma = O(1)$, la borne de complexité est $O(\sqrt{n} L)$. ■

Conclusion

Dans ce travail, nous avons étudié essentiellement deux méthodes de points intérieurs pour résoudre le problème de programmation linéaire (PL).

Ces deux méthodes sont basées sur la résolution de problèmes intermédiaires qui diffèrent en leur fonction objectif :

La fonction de pénalité intérieure f_α pour l'un et la fonction potentielle f_q pour l'autre.

La résolution exacte des problèmes intermédiaires génère des suites de points centraux $(x(\alpha))$ et $(x(v))$. Ces suites décrivent la même trajectoire centrale et convergent vers une solution optimale du problème (PL).

Des algorithmes polynomiaux sont obtenus en suivant cette trajectoire de manière approximative, selon un critère de proximité et ont une complexité de l'ordre de $O(nL)$ et $O(\sqrt{n}L)$ itérations.

Un complément intéressant de ce travail serait de développer les procédures d'initialisation et de "purification" [8] qui ne furent qu'énoncées ici même.

Bibliographie

- [1] C.GONZAGA, Polynomial affine algorithms for linear programming. Internal report ES-141/88, COPPE-Federal University of Rio de Janeiro, Brasil, 1988; Math.Programming, 49, p 7-21,(1990).
- [2] C.GONZAGA, Large step path-following methods for linear programming. part I : Barrier function method. Internal report, COPPE-Federal University of Rio de Janeiro, Brasil, 1989; SIAM J.Optimization vol.1, p 268-279, (1991).
- [3] C.GONZAGA, Large step path-following methods for linear programming. part II : Potential reduction method. Internal report, COPPE-Federal University of Rio de Janeiro, Brasil, 1989; Siam J.Optimization vol.1, p 280-292, (1991).
- [4] C.GONZAGA, Path following methods for linear programming. Publication technique, COPPE/UFRJ, Brasil, 1989.
- [5] M.GROTSCHER, L.LOVASZ and A.SCRIJVER, Geometric algorithms and combinatorial optimization. Springer Verlag, Berlin (1988), p 21-35.
- [6] H.IMAI, On the convexity of the multiplicative version of Karmarkar's potential function. Math.Programming, 40 (1988), p 29-32.
- [7] N.KARMARKAR, A new polynomial time algorithm for linear programming. Combinatorica, 4 (1984), p 373-395.
- [8] K.O. KORTANEK and ZHU JISHAN, New purification algorithms for linear programming. Department of management science, University of Iowa, Iowa city. Naval Research logistics, vol.35, p 571-583 (1988).
- [9] J.M. ORTEGA and W.C. RHEINBOLDT, Iterative solution of linear equations in several variables. Computer Science and applied mathematics, Academic Press (1970).
- [10] C.H. PAPADIMITRIOU and K. STEIGLITZ, Combinatorial optimization : Algorithms and Complexity. Prentice-Hall, (1982).
- [11] POLAK, Computational methods in optimization. Mathematics in science and engineering, vol. 77, Academic Press New-York and London, (1971).